Universidad Autónoma de Madrid
Facultad de Ciencias, Departamento de Biología Molecular

DIRECTED EVOLUTION OF ANCESTRAL AND MODERN ENZYMES

Bernardo José Gómez Fernández

TESIS DOCTORAL

Madrid, 2019

Universidad Autónoma de Madrid
Facultad de Ciencias
Departamento de Biología Molecular

Bernardo José Gómez Fernández
Licenciado en Biotecnología

Director
Dr. Miguel Alcalde Galeote
Instituto de Catálisis y Petroleoquímica
Consejo Superior de Investigaciones Cientificas

Tutora
Dra. Elena Bogónez Peláez
Universidad Autonoma de Madrid
Facultad de Ciencias, Departamento de Biología Molecular

Madrid, 2019

AGRADECIMIENTOS/ACKNOWLEDGEMENTS

Quiero agradecer a todas y cada una de las personas que me han apoyado durante todos estos años, que han ayudado a que el desarrollo de esta Tesis haya sido una gran experiencia a nivel profesional y personal, gracias.

En primer lugar agradecer a mi Director de Tesis, el Dr. Miguel Alcalde, primero por darme la oportunidad de entrar en su laboratorio y posteriormente por confiar en mí para abordar proyectos de diferente índole. Valoró especialmente que me hay brindado la oportunidad de realizar varias estancias durante todos estos años, una experiencia enriquecedora a todos los niveles. Su entusiasmo por lo que hacemos, la dedicación e interés en ayudarme a mejorar profesionalmente, y lo más importante, su trato cercano (hasta el punto de intercambiarnos *comics*), han contribuido enormemente a que haya disfrutado de esta Tesis en todos los aspectos. Gracias también por confiar de nuevo en mí para formar parte de nuestra nueva aventura, EvoEnzyme.

A mi profesor de cuarto curso de carrera, el profesor Plou, (Kiko). Con el descubrí lo que la evolución dirigida podía hacer y además me puso tras la pista de Miguel. Seguramente el camino a seguir hubiera sido diferente sin estas casualidades. Quiero agradecerle también por todo lo que ha aportado en gran parte de mi Tesis. Sin mencionar lo que he disfrutado viéndole actuar sobre el escenario. Al Dr. Antonio Ballesteros, nadie mejor que uno de los padres de la biocatálisis en España para llevar a cabo una Tesis con enzimas, gracias por tus consejos y el entusiasmo con el que nos alientas, es inestimable. A mi tutora académica, la Dra. Elena Bogónez. Por su eficiencia y simpatía a la hora de llevar a cabo las gestiones propias de la Tesis y por sus valiosas palabras de ánimo durante estos años.

La presente Tesis Doctoral se ha llevado a cabo gracias a la financiación recibida a través de una beca para la formación de personal investigador (FPI) del Ministerio de Economía y Competitividad (BES-2014-068887) dentro de los proyectos nacionales DEWRY (BIO2013-43407-R) y LIGNOLUTION (BIO2016-79106-R). Especial agradecimiento al CTQ de REPSOL S.A, por su apoyo y financiación a través de los proyectos Rubolution y Rubolution 2.0., fundamentales para la evolución de Rubiscos modernas y ancestrales; y de manera muy especial a Mónica García Ruiz y María Luisa Rodríguez Buey del CTQ por su apoyo e ideas constantes durante esta primera etapa. También doy las gracias a la Acción Cost- CM1303: System Biocatalysis y al proyecto

Proyecto PIE-CSIC 201780E043 “Evolución dirigida de rubisco para modificar el balance carboxilasa/oxigenasa”.

Del CIB quiero agradecer especialmente a las Dras. Alicia Prieto y María Jesús Martínez. Gracias a ellas di el paso final para empezar mi Tesis con Miguel. También les agradezco enormemente el trato y la dirección recibida por su parte durante mis prácticas en su laboratorio. También quiero expresar mis agradecimientos a la Dra. Susana Camarero por sus consejos referentes a lacasas y a los Dres. Ángel Martínez y Javier Ruiz Dueñas. De mis compañeros especial agradecimiento a Iván Ayuso, a nivel profesional por todo lo compartido respecto a las enzimas ancestrales. A nivel personal porque es un tipo estupendo con mucha paciencia para quedar conmigo. A Manolito, una persona con la que me siento muy cómodo y tuve la suerte de conocer al principio de todo esto, comienzos así hacen que el camino valga la pena. También quiero agradecer a Felipe (eres un crack), Julia, Lucia, Norge, Juan Carro, Rosa, Isa P., Carlos, Lidem, Lola, Mariu, Rosa, María, Vero, Davi y el resto de gente con la que pasé esos meses en el CIB y posteriormente en cañas y casas rurales.

De mi estancia en la Facultad de Ciencias de Granada, agradezco especialmente al Prof. José Manuel Sánchez Ruiz por darme la oportunidad de acudir a su laboratorio a aprender sobre la reconstrucción y resurrección de enzimas ancestrales. Toda la temática de resurrección ancestral de la Tesis se basa en lo que aprendí en su laboratorio por lo que ha sido de capital importancia para el buen desarrollo de esta Tesis. A la Dra. Valeria Risso del grupo del Prof. Sánchez Ruiz, que tras su paso por nuestro laboratorio me devolvió con creces la hospitalidad recibida. Gracias también por toda la ayuda que me ha brindado posteriormente, ha sido un placer colaborar contigo. También quiero agradecer al Dr. Ángel L. Pey que me ayudo con los experimentos de ITC y fluorimetría, por su paciencia y dedicación. Quiero agradecer también a la gente con la que compartí buenos momentos en esa estancia, como Adela, Encarni, Fadia, Sergio, Asun y el resto de compañeros.

I want to thank Dr. Stefan Lutz from University of Emory in Atlanta, Georgia, USA, for giving me the opportunity to spend three months in his laboratory where I learnt the basis on circular permutation, while developing my skills in the use of *Pichia pastoris*. Thanks to Sam, Leann and Matt for their assistance and patience, as well as for the funny moments and jokes (and weird videos) of everyday. I would also like to thank Michelle, Parisa, Angel, Huanyu, Elsie, Evy and David who contributed to make my stay a wonderful experience. Por supuesto, agradezco a mi familia americana, Carol, Renato y Benji por lo bien me acogieron en su casa como uno más de la familia, os tengo siempre presentes.

Special thanks to Dr. Spencer Whitney from the Australian National University (ANU) in Canberra, Australia, for giving me the opportunity to learn about rubisco engineering from one the biggest experts in the world. In addition, I want to give him my most sincere gratitude for his kindness, taking me to uncountable activities and sport events, truly far beyond of what it is expected from a supervisor during a stay (even singing in the lab…), thank you!. Thanks also to Manny, Ding and Amy for all the tours, beers and great time together. Also thanks to Rosemary, Brendon, Rob, Sophie, Tim and Sally.

Como no podía ser de otro modo, agradezco especialmente a mis compañeros de laboratorio del ICP por varias razones. A Eva que fue mi supervisora en mis inicios, por el cariño y la dedicación con la que me ayudo a comenzar, es un placer volver a tenerte por aquí. A Diana por todos los consejos y ayudas con respecto a la lacasas, también por las risas y las charlas que hemos tenido. A David además de por la ayuda prestada en el laboratorio, por todos esos momentos tan divertidos en las diferentes cafeterías de la UAM donde llegamos a desarrollar interesantes hipótesis sobre monitos y la luna. A la Wiser, porque eres más maja que las pesetas y aún más salada, tu venida al ICP nos dio mucha alegría a todos. A Javi, mi compañero de aventuras al inicio de todo esto con el que además he compartido risas y alguna que otra escapada. A la Patch (yo lo escribo así) porque gracias a ella dosifico el número de veces que digo "Patrrricia" al día y así tengo más saliva para hablar de otras cosas, además de suponerme un buen apoyo cuando necesitaba quejarme un poco. A Iván, el "birdo" al que aprecio por muchas razones, desde nuestras interesantes charlas sobre lacasas hasta nuestro gusto mutuo por la buena comida acompañada de buen vino. A Xavi "cracko", porque "Frodo no habría ido tan lejos sin Sam", no sé si sabes lo mucho que significan estas palabras para mí, te las dedico de corazón, gracias. Agradezco también a los estudiantes que he tenido bajo mi supervisión. Especialmente a Leyre, que además de aprender ella de mí y yo de ella, hemos sacado una amistad maravillosa, para cuidarla mucho tiempo. También a Berndjan y Andrés, con los que disfruté y aprendí cosas que no siempre tienes la suerte de encontrar durante la Tesis. Por último a todas las personas que han pasado por el laboratorio a lo largo de estos años: Dat, Gordana, Morgane, Elvin, Sofía, Kat, Jia, Juan, Mehdi, Joaquín, Mar, Alejandra, Diego y más. Del laboratorio de Kiko, a Paloma, por todo ese tiempo trabajando en el "Rubis team" así como por el tiempo disfrutado en festivales o digo más, en la Syfy. A Noa, que poco a poco me admite dentro del grupo selecto de gallegos no-gallegos, gracias. A Fafa, pieza fundamental en las risas de la comida y con la que tengo charlas muy interesantes. A Joselu que me mantiene al día con la actualidad política. A David, porque da gusto pasar el

rato con él, sobre todo si es de fiesta. A Lu y Barbis, por ser tan cariñosas y divertidas. Del resto del ICP quiero hacer un agradecimiento general a todas las personas con las que por unas razones u otras he interactuado a lo largo de estos años, especialmente a Lara, Moni, Coscolín, David, Chiara, Cristina, Oscar, Alejandro, Javier, María, Rafa, Elena, Janaina y tantos otros más. También agradezco al Dr. Manuel Ferrer por los equipos prestados y a la Dra. Consuelo Álvarez, nunca una mejor compañera de oficina. También agradezco a todo el personal de administración, de limpieza, de mantenimiento y de informática, entre todos hacemos que este edificio tan feo sea un lugar acogedor gracias a las personas que lo forman.

A esos amigos que sin estar involucrados en el laboratorio o el proyecto me han ayudado a lo largo de estos años. A mi amigo y compañero de piso Chusmi, por todas las horas que dedicamos juntos a sacar adelante nuestras Tesis cuando el día ya había terminado, así como todas las risas y los muy buenos momentos que hemos pasado juntos. A Hernán, que además de descubrirme las clases de aerobic de Carmen, me echo un cable más de un vez cuando estabas atascado en algún experimento. A Iñigo, que por el módico precio de un par de cervezas siempre ha estado más que dispuesto a ayudarme con la parte computacional.

También quiero agradecer a todos los amigos y familia que han estado ahí para hacer que los momentos de ocio fueran increíbles. Me gustaría poneros a cada uno una dedicatoria pero esto ya se está alargando.

A la pandilla "Frikis" con los que paso la mayor parte del tiempo en Madrid, sois un conjunto de personas maravillosas de los que me siento muy orgulloso y afortunado de formar parte. También a Nacho, Charlie, Carlos, Canario y otros amigos con los que paso tiempo. Por supuesto no me olvido de antiguos compañeros de la carrera como Manu, Laura, Mer, Zule, María y más. También a mis compañeros del universo paralelo en la Complu Diego, Dani, Jako, Marta, Laura y Tapetado. Por ultimo a mis actuales compañeros de piso actuales Edu, Miguel y Nafri, que me han tenido que sufrir durante los meses de la escritura.

A Zamora, mi ciudad, mi hogar, mi casa. A todos mis amigos de "Hienas salvajes" tanto por los momentos de fiesta y diversión así como los cafés, momentos de frikeo, vacaciones, nocheviejas, cumpleaños y todo lo que conlleva en algunos casos más de 20 años de amistad. Creo que hasta ahora hemos cumplido muchas de las locuras que nos planteamos en su día y eso es decir mucho. También quiero agradecer a mis antiguas

compañeras del "Cora" con la que sigo en contacto Marina, Yaiza y Silvia, por los momentos tan agradables que seguimos pasando juntos.

A mi familia, a todos ellos. Tengo la enorme suerte de tener una muy numerosa y con muy buena composición. Empezando por mis primos con los que mantengo una muy buena relación, mención especial a Pablo que en cierto modo algo ha tenido que ver en algunas de mis decisiones profesionales. De mis tíos y tías, destaco a mi padrino Pedro (el mus) y mi madrina Pepita (eres todo amor). También destaco a mis "madrinas madrileñas" Mabel, por la comida y el enorme cariño que me has dado y a Carmina la otra Doctora de la familia y mi madrina de la carrera, tu constante apoyo y dedicación para el bien hacer de mi carrera y Tesis han sido claves para llevarlas a buen puerto, que no te quepa duda. A mis cuatro abuelos, que aunque no están con nosotros, las raíces son la base y está familia tiene los mejores cimientos, os quiero haya donde estéis. Por supuesto agradecimiento a los nuevos integrantes Martín e Isabel, ¡sois los tesoros de la familia! También a mi familia Segoviana que siempre me acogen con mucho cariño.

A Mari, mi segunda madre. El tenerte en casa cuando voy a Zamora es un lujo y ya no digamos disfrutar de tu cocina y tu cariño.

A ti, por todo. Cualquier cosa que escriba aquí no haría justicia a lo que siento.

A mi hermana, mi ojito derecho. Siempre estás en mi cabeza y cada día descubro nuevos niveles de orgullo de hermana pequeña, compartir ADN contigo es una pasada, te quiero. Con Tere viene en pack el agradecer a Víctor (por cuidar tan bien de ella) y a Bartolo, que aunque él no se dé cuenta ha contribuido enormemente a hacer que los momentos malos se volviesen buenos y alegres, no nos merecemos a los perros.

A mis padres, Bernardo y Magdalena. Esta tesis es tanto mía como vuestra. Os la dedico porque sin vosotros en la ecuación no habría pasado, yo no sería quien soy y soy una persona muy afortunada y feliz. La gente de la que me rodeo, todos los amigos y familiares que están aquí mencionados son, en gran medida, el resultado de que unas personas buenas, divertidas y sensatas me han brindado lecciones de vida de primera. Gracias y perdón, porque con una vida no me basta para devolveros todo lo que me dais.

SUMMARY

Ancestral sequence reconstruction (ASR) and resurrection (*i.e.* functional expression in a heterologous host) allows enzymes with different properties to be disclosed while its combination with directed evolution may lead to the development of a new generation of biocatalysts. In this Doctoral Thesis we have explored the combination of such powerful methods using as blueprints two different enzyme systems, Rubisco and laccase.

In the first chapter of this Thesis we reconstructed and resurrected (in *Escherichia coli*) Precambrian Rubisco nodes which were evolved in parallel with the extant Rubisco counterpart. An *in vitro* dual high-throughput screening (HTS) method was set out to identify thermostable and functional variants after- applying a palette of directed evolution strategies. The stronger tolerance to mutational loads, the improved expression and the different kinetic behavior were some of the traits that highlighted in the Precambrian enzyme. Particularly, the evolved ancestral Clone B2 stood out as a case study of this elusive protein due to its alternative performance in the classical equilibrium of Rubisco kinetic constants.

In the second chapter we focused ASR and directed evolution on basidiomycete laccases. Firstly, ancestral nodes from the Paleozoic were reconstructed and resurrected in *Saccharomyces cerevisiae*. The resurrected enzymes showed a higher heterologous expression and a broader pH stability profile than the modern -laboratory evolved- counterpart. The most promising ancestral node was subjected to structure-guided evolution for the oxidation of β–diketones, an unusual type of redox mediators capable to initiate the polymerization of vinyl monomers.

The final chapter of the Thesis deals with consensus design, a long-standing protein engineering method to increase stability without compromising activity. We applied an in-house consensus method to stabilize a laboratory evolved high-redox potential laccase. Multiple sequence alignments were carried out and computationally refined by applying relative entropy and mutual information thresholds. Through this approach, an ensemble of 20 consensus mutations were identified, 18 of which were consensus ancestral mutations. After analyzing potential epistasis by site directed recombination *in vivo*, the best mutant was characterized displaying dramatically improved thermostability, kinetic values and secretion levels.

RESUMEN

La reconstrucción y resurrección (*i.e.* expresión funcional en un hospedador heterólogo) de secuencias ancestrales permite obtener enzimas con diferentes propiedades que al ser combinadas con la evolución dirigida pueden dar lugar al desarrollo de una nueva generación de biocatalizadores. En la presente Tesis Doctoral hemos explorado la combinación de estos potentes métodos usando como modelos dos sistemas enzimáticos diferentes, Rubisco y lacasa.

En el primer capítulo se reconstruyeron y resucitaron (en *Escherichia coli*) nodos de rubiscos Precámbricas con el fin de evolucionarlos en paralelo con una versión moderna de Rubisco. Se desarrolló un método de cribado dual *in vitro* para poder identificar variantes termoestables y funcionales tras aplicar varias estrategias de evolución dirigida. Las enzimas Precámbricas destacaron por una alta tolerancia a tasas mutagénicas, expresión funcional mejorada y valores cinéticos diferentes a los de las enzimas modernas. En particular, la rubisco ancestral evolucionada, clon B2, despuntó como caso de estudio de esta complicada enzima debido al comportamiento alternativo que muestra con respecto al equilibrio clásico de las constantes cinéticas de la Rubisco.

En el segundo capítulo se llevo a cabo la resurrección y evolución dirigida de lacasas de basidiomicetos. En primer lugar se reconstruyeron y resucitaron en *Saccharomyces cerevisiae* nodos ancestrales del Paleozoico. Las enzimas ancestrales mostraron mayor nivel de expresión heteróloga así como un perfil de estabilidad a diferentes pHs más amplio que el de la versión –evolucionada en el laboratorio- moderna. El nodo ancestral más prometedor fue sometido a evolución estructuralmente guiada para la oxidación de β-dicetonas, un tipo de mediador redox poco usual capaz de iniciar la polimerización de monómeros de vinilo.

El capítulo final de la Tesis trata sobre el diseño consenso, un método clásico de ingeniería de proteínas para aumentar la estabilidad sin afectar a la actividad. Se aplicó un método consenso propio para la estabilización de una lacasa de alto potencial redox evolucionada en el laboratorio. Se llevó a cabo un alineamiento de múltiples secuencias que fue refinado computacionalmente mediante el uso de los marcadores de entropía relativa e información mutua. Mediante este procedimiento se identificaron 20 mutaciones consenso, 18 de las cuales corresponden a mutaciones ancestrales-consenso. Se analizó la posible

epistasia de estas mutaciones mediante recombinación dirigida *in vivo* y se caracterizó el mejor mutante que presentó mayores niveles de estabilidad, valores cinéticos y secreción.

CONTENTS

ACRONYMS

2-PG	2-phosphoglycolate
3-PGA	3-phosphoglycerate
AAO	Aryl-alcohol oxidase
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
AMM	Ancestral Mutation Method
Amp	Ampicillin
ASR	Ancestral Sequence Reconstruction
ATP	Adenosine Triphosphate
BLASTp	Basic Local Alignment Search Tool for proteins
BPGA	1,3-bisphosphoglycerate
bp	Base pairs
BSA	Bovine serum albumin
C-CABP	2-carboxyarabinitol-1,5-bisphosphate
CSM	Combinatorial saturation mutagenesis
CV	Coefficient of variation
D1, D2, D3	Cupredoxin domains 1, 2, 3
DHAP	Dihydroxyacetone-P
DMP	2,6-Dimethoxyphenol
DMSO	Dimethyl sulfoxide
dNTPs	Deoxyribonucleotides
DOGS	Degenerate oligonucleotide gene shuffling
DSC	Differential scanning calorimetric
E	Redox potential
ECM	Enzyme-CO_2-Mg^{2+}
EDTA	Ethylenediaminetetraacetic acid
ELSD	Evaporative light scattering detector
E_{T1}	Redox potential at T1 site
EtOH	Ethanol
epPCR	Error-prone PCR
ET	Electron transfer
FACS	Fluorescence-activated cell sorter

FPLC	Fast protein liquid chromatography
FRESCO	Framework for Rapid Enzyme Stabilization by Computational libraries
G3P	Glycerol-3P
GAP	Glyceraldehyde-3P
GAPDH	Glyceraldehyde-3P dehydrogenase
GARLI	Genetic Algorithm for Rapid Likelihood Inference
GPDH	Glycerol-3P dehydrogenase
Gyr	Gigayear
HAT	Hydrogen-atom transfer
HBT	1-hydroxybenzotriazole
HPLC	High performance liquid chromatography
HRPLs	High redox potential laccases
HTS	High throughput screening
IA	Initial activity
IPTG	Isopropyl ß-D-1-thiogalactopyranoside
ITCHY	Incremental Truncation for the Creation of Hybrid enzYmes.
IVOE	*In vivo* Overlap Extension
JTT	Jones–Taylor–Thornton. Jones substitution model
kb	Kilobase
K_C	Rubisco affinity for CO_2 under no O_2 atmosphere at 25ºC
k_{cat}	Catalytic constant
k_{cat}^{C}	Catalytic constant for rates of CO_2-fixation
k_{cat}/K_m	Catalytic efficiency
k_{cat}^{C}/K_C	Carboxylation efficiency
K_m	Michaelis-Menten constant
L	Large subunit
LacAnc95	Laccase ancestor node 95
LacAnc98	Laccase ancestor node 98
LacAnc100	Laccase ancestor node 100
LB	Luria-Bertani
LccED	Laccase and Multicopper Oxidase Engineering Database
LMS	Laccase mediator system
MI	Mutual information

ML	Maximum likelihood
MSA	Multiple sequence alignment
MORPHING	Mutagenic Organized Recombination Process by Homologous *In vivo* Grouping
MRβ/γPro	The most recent common ancestor of β and γ proteobacteria
MRPro	The most recent common ancestor of all the proteobacteria
MRProFir	The most recent common ancestor of proteobacteria and the cyanobacteria firmicutes
MUSCLE	MUltiple Sequence Comparison by Log-Expectation
MW	Molecular weight
NADPH	Nicotinamide adenine dinucleotide phosphate
NCBI	National Center for Biotechnology Information
NHE	Normal hydrogen electrode
OD_{600}	Optical density at 600 nm
PAML	Phylogenetic Analysis by Maximum Likelihood
PAUP	Phylogenetic Analysis Using Parsimony
PcL	*Pycnoporus cinnabarinus* laccase
PCR	Polymerase chain reaction
PDA	Photodiode array
PDB	Protein data bank
PFAM	Protein Families
PGK	Phosphoglycerokinase
PM1L	PM1 laccase
PNGase	Peptide N-glycanase
PROSS	Protein Repair One Stop Shop
RA	Residual activity
RACHITT	Shuffling towards highly evolved genes
RASPP	Recombination as Shortest Path Problem
RDE	Rubisco dependent *Escherichia coli*
RE	Relative entropy
RLP	*Rubisco like protein*
RuBP	Ribulose-1,5-bisphosphate
RubRr	Rubisco form *Rhodospirillum rubrum*
SDR	Side Directed Recombination

SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEM	Selective expression medium
SHIPREC	Sequence homology-independent protein recombination
S	Small subunit
$\mathbf{S_{C/O}}$	Specificity for CO_2 over O_2
SMART	Simple Modular Architecture Research Tool
StEP	Staggered extension process
$t_{1/2}$	Half life. Time required by the enzyme to lose 50% of its initial activity upon incubation at a given temperature
T_{50}	The temperature where the enzyme lost 50% of its initial activity, after 10 minutes of incubation
T1	Type 1 copper
T2/T3	Trinuclear Cu cluster
TAI	Total activity improvement
TEMPO	(2,2,6,6-Tetramethylpiperidin-1-yl)oxyl
TPI	Triose-P isomerase
Tsp	Total soluble protein
V_C^{max}	Maximal rate of carboxylation
VLA	Violuric acid
YNB	Yeast nitrogen base
α^{native}	Native α factor prepro-leader
α^{PcL}	Evolved α factor prepro-leader for PcL secretion
α^{PM1L}	Evolved α factor prepro-leader for PM1L secretion
Ɛ	Extinction molar coefficient

1 INTRODUCTION

1.1 Enzymes within the framework of directed evolution

Enzymes are biocatalysts that hasten chemical reactions in all living organisms. Directly extracted from the vast natural diversity, where they performed specific biological functions, enzymes are employed as environmentally friendly `molecular workhorses´ to conform a broad portfolio of practical biotechnological solutions made available to the humankind. While enzymes are promiscuous (in terms of substrate but also catalytic promiscuity), they also show exquisite stereo-, regio- and chemo-selectivity with the added advantage of displaying activity in conditions that fulfil the green standards (Molina-Espeja et al. 2016; Nestl et al. 2011; Alcalde et al. 2006). Sadly, when an enzyme is included within a harsh industrial process, it cannot withstand the specific reaction requirements, thereby ending in most of the cases inactivated: After all, natural enzymes are simply the product of thousands of years of natural molecular evolution and therefore their functions are the direct consequence of the conditions marked by the time/place in which they have been enclosed and driven by the role they play within a given ecosystem. To adapt natural enzymes to the industry, protein engineering is the strategy of choice (Narancic et al. 2015; Bornscheuer and Buchholz 2005; Kuchner and Arnold 1997; Skandalis et al. 1997), and above all, directed evolution has stood out as a disruptive and straightforward technology. Recently recognized with the Nobel Prize in Chemistry 2018, `the directed evolution decalogue´ established by Frances Arnold from the California Institute of Technology is based on molecular breeding and exploration of genetic diversity, harnessing different methods of random mutagenesis, DNA recombination and artificial screening/selection. Suitable heterologous hosts are needed in order to express the corresponding enzyme mutant libraries from that genetic pool. Afterwards the best clones are screened under the selective pressure established by the researcher with the aim of finding improved candidates that are included in new rounds of directed evolution, until the desired biochemical trait arises (Arnold 2018; Alcalde 2017a; Cobb et al. 2013) **Figure 1.1.1**.

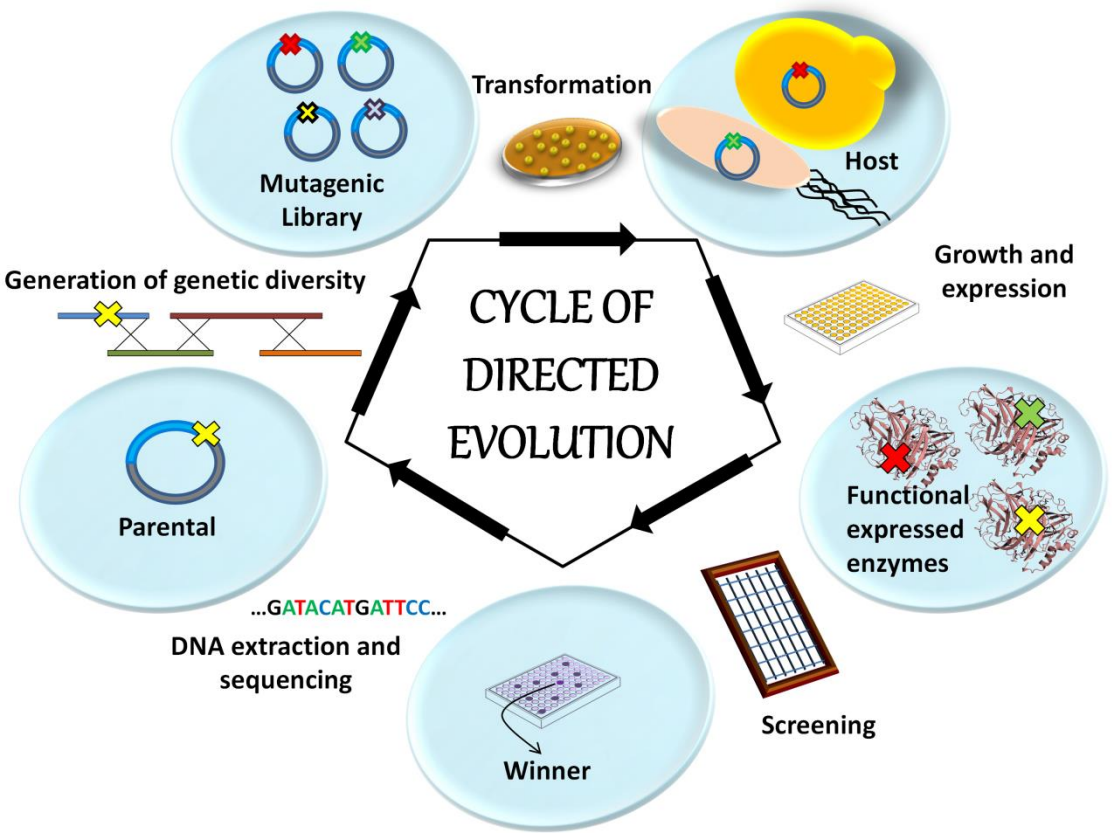


Figure 1.1.1 Representation of a standard directed evolution cycle including: Generation of genetic diversity, transformation of the mutant library in an appropriate host (typically yeast or bacteria), expression of the mutant library, high-throughput screening and selection of the mutant `winner´ to be used as parental for a new evolution round. Mutations are represented as crosses.

Accordingly, high-throughput screening (HTS) assays or biological selection systems, DNA library creation methods and suitable hosts for functional expression (mostly *Escherichia coli* or *Saccharomyces cerevisiae*) are the three critical components in a successful directed evolution campaign. In terms of screening mutant libraries, it is critical to have a rapid, sensitive and reliable detection method of the activity or property intended to improve, preferably associated to colorimetric or fluorimetric responses. The incorporation of flow-citometry systems (*e.g.* Fluorescence-activated cell sorter (FACS)) along with the use of *in vitro* or *in vivo* compartimentalization is giving a new twist in the development of sophisticate ultra HTS protocols (Mate et al. 2017). Conversely, library creation depends strongly on the chosen source for genetic diversity (Packer and Liu 2015; Shivange et al. 2009; Reetz et al. 2008). The classical method to introduce variability is random mutagenesis. This can be achieved through different strategies such as i) error prone polymerase chain reaction (*ep*PCR) of the whole protein template (replacing randomly single positions by arbitrary substitutions, -although the method is far to be

perfect as it is biased to certain substitutions linked to the chosen DNA polymerase and/or specific mutagenic conditions), ii) Focused –(structure-guided) mutagenesis in defined protein regions -on the base of previous structural analysis in combination with computational simulations- through, for instance saturation mutagenesis whereby single residues are replaced by the 20-naturally occurring amino acids to explore the whole protein alphabet in such defined position/s (Gonzalez-Perez et al. 2014; Reetz and Carballeira 2007; Chica et al. 2005; Williams et al. 2004; Georgescu et al. 2003; Cline and Hogrefe 2000; Cline et al. 1996). Besides classical directed evolution (*i.e.* based on the selection of the fittest), neutral genetic drift is arising as an invaluable approach aimed at opening new adaptive pathways by purging deleterious mutations while creating a polymorphic population of neutral variants, with enhanced promiscuity and stability (Gupta and Tawfik 2008; Bloom et al. 2007).

Although all these tools have proved to be successful in directed enzyme evolution, they are lacking of the best pool of genetic diversity, the one recorded in nature through hundreds of years of natural evolution based on sexual reproduction and fitness. As such, a number of techniques are also available for researches to further exploit such information for library design (Cole and Gaucher 2011). Indeed, there are tens of DNA recombination methods (Family shuffling, Staggered extension process (StEP), Shuffling towards highly evolved genes (RACHITT), Degenerate oligonucleotide gene shuffling (DOGS), Incremental Truncation for the Creation of Hybrid enzYmes (ITCHY), Sequence homology-independent protein recombination (SHIPREC), SCHEMA etc.), all of them imbibe from the whole sequence information enclosed in modern proteins, recombining enzymes to generate a library of variants with hybrid properties (Mate et al. 2017 and references herein; Kikuchi and Harayama 2002; Ness et al. 2002; Keightley and Eyre-Walker 2000; Stemmer 1994). Besides classic experimental recombination protocols, computational methods based on Rosseta (*e.g.* Protein Repair One Stop Shop (Pross), FunLib) (Khersonsky et al. 2018; Goldenzweig et al. 2016) are becoming important within the protein engineering toolbox.

Particularly, some of them are uniquely related with multiple sequence alignment (MSA), like consensus design and enzyme resurrection, and as such they can help reduce the exploration of vast protein sequence space, to enhance stability, promiscuity and unmask other relevant biochemical characteristics. However, while consensus design takes efficiently the information of modern enzymes, ancestral sequence reconstruction (ASR) gives researchers access to the genetic data by traveling back in the timeline of evolution,

opening new possibilities in library design (Risso et al. 2018; Alcalde 2017b; Merkl and Sterner 2016a; Cox and Gaucher 2014;), **Figure 1.1.2**.

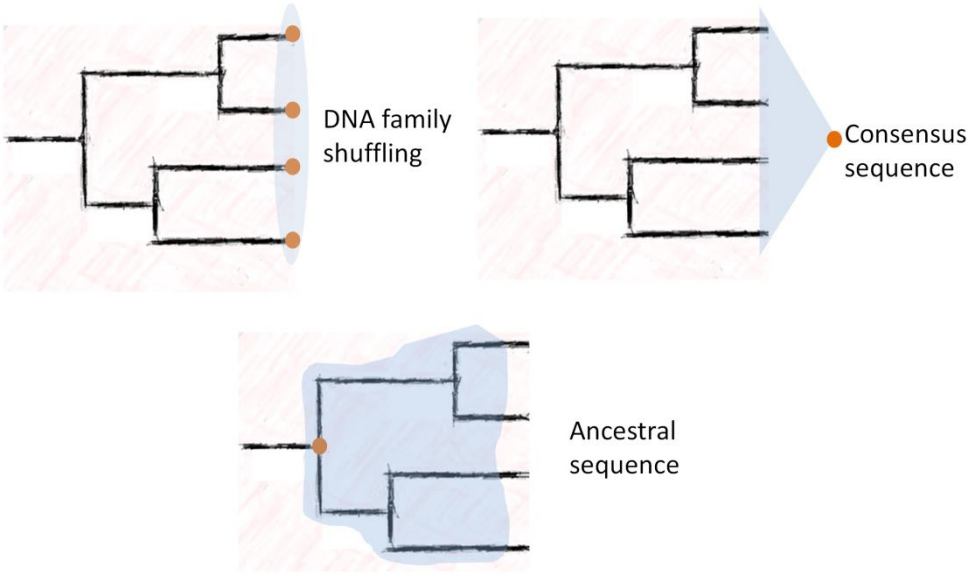


Figure 1.1.2 Representation of different strategies involved in the exploration of the natural genetic diversity. DNA recombination (*e.g* by DNA family shuffling) is conceived to explore the information encoded in modern orthologues; consensus design extracts the genetic diversity from the same pool as DNA family shuffling but focusing on single specific residues to yield a single consensus sequence; ASR explore the information enclosed through years of natural evolution to infer the sequences of the inner nodes of the phylogeny.

In this Doctoral Thesis, we have paid special attention to the use of consensus mutagenesis and ASR for the generation of genetic diversity within directed evolution campaigns with two different enzymes, both involved in the carbon cycle in biosphere: Rubisco and laccase. Accordingly, the next sections are including general descriptions of these strategies followed by the introduction of both enzyme systems.

1.1.1 Consensus Method

The distribution of amino acids extracted from MSA of homologous proteins reveals the most conserved positions along the different sequences (Lehmann and Wyss, 2001; Lehmann et al. 2000). The majority of these residues are considered to be stabilizing since they can help withstand the demanding adaptation process that takes place in the course of natural evolution. The identification and insertion of such consensus mutations in a protein template is known as consensus design and it contributes to stabilize enzymes while bypassing the unwanted activity/stability tradeoff (Siddiqui 2017; Porebski and Buckle 2016; Lehmann et al. 2002). Comparing with other long-standing strategies to

improve thermostability (*e.g.* classical directed evolution, rational rigidification of flexible loops, computational methods like FRESCO, SCHEMA or PROSS, ancestral libraries and protein resurrection) (Reetz 2017; Magliery 2015; Nestl and Hauer, 2014; Cole and Gaucher, 2011), consensus design encloses some appealing advantages that make it a good choice to be applied individually or in combination with the abovementioned methods. Indeed, consensus mutagenesis finds an optimum balance between library-size and the required knowledge of the protein, which allows smart and functionally-enriched mutant libraries to be generated from MSA while streamlining resources and research effort (Jackel et al. 2010; Lehmann et al. 2002; Lehmann et al. 2000).

The concept of consensus design was applied for the first time in 1989 by Pantoliano and coworkers who improved the dynamic stability of a protease. But it was not until 1994 when Steipe, through a study based on an immunoglobulin domain, showed its true efficiency in protein engineering (Steipe et al. 1994; Pantoliano et al. 1989), **Figure 1.1.3**. The consensus approach establishes that all the amino acids of a given protein contribute in some extent to the overall stability and it is typically carried out according to the following steps:

i) Search of homologous sequences. This can be done by using databases like National Center for Biotechnology Information (NCBI) (Coordinators 2018), Swiss-Prot (UniProt 2019), Protein Data Bank (PDB) or from more specialized sources of information such as Protein Families (Pfam) (Finn et al. 2016; Sonnhammer et al. 1998), Simple Modular Architecture Research Tool (SMART) (Letunic et al. 2015), Prosite (Sigrist et al. 2013) or Superfamily (Wilson et al. 2009). NCBI and PDB usually render bigger number of sequences than Pfam, SMART Prosite or Superfamily, whose information has been previously curated. In this regard, it is worth noting that reasonable consensus designs can be achieved with less than 20 sequences (Jones et al. 2017; Jäckel et al. 2010; Lehmann et al. 2002), yet by decreasing the diversity of sequences, the signal-to-noise ratio and the statistical significance are reduced (Porebski and Buckle 2016).

ii) Elaboration and curation of the MSA. The standard tools for MSA are Clustal X, Clustal W (Larkin et al. 2007) and MUltiple Sequence Comparison by Log-Expectation (MUSCLE) (Edgar 2004). Some hurdles arise during MSA (*i.e.* sequences too short or too large, repeated sequences and deletions/insertions), and they must be curated. As a rule of thumb, sequences too short are discarded. Longer sequences than the average may be reduced before their final inclusion. It is

also important to notice that databases tend to generate bias during consensus design as the sequencing of some genomes is more demanded than others, thereby enriching the number of sequences available from certain organisms (Jäckel et al. 2010). This problem may be circumvented by deleting repeated sequences or by avoiding in the query those sequences with a high degree of identity between them (Lee 2014). Concerning deletions and insertions, it is recommended to calculate their degree of occupancy for a giving position, keeping the ones that present a percentage of occupancy above the established threshold (Pearson 2013; Sullivan et al. 2011; Roost 1999).

iii) Calculation of Consensus residues: The most repeated residue for a given position is referred to as consensus. Some metrics may help to sharp the query upon finding those conserved residues with biological significance. For example, Relative Entropy (RE) takes into account a reference distribution which typically is based on the codon usage within the yeast proteome (Sullivan et al. 2012; Magliery and Regan 2005). This distribution is bound to physiological laws and therefore it can be used as reference for determining those positions that are highly conserved.

Certainly, there are some bioinformatics tools that implement the above-mentioned steps in a few clicks, yet the user has less control over the selection than when is manually done (Durani and Magliery 2013; Jones et al. 2017; Musil et al. 2017; Goldenzweig et al. 2016). Once the consensus residues are elucidated they can be filtered via different criteria (Sullivan et al. 2012; Magliery and Regan 2005; Applebaum 1996), and then evaluated by mutagenesis on the query template (Polizzi et al. 2006; Wang et al. 1999; Nikolova et al. 1998). These consensus mutations can be added to directed evolution schemes (for instance by DNA shuffling) (Case and Hackel 2016; Khersonsky et al. 2012; Bershtein et al. 2008); a rougher yet faster strategy is the synthesis of the consensus enzyme (*i.e.* including all consensus mutations, without sieving the residues) albeit at the cost of jeopardizing catalytic activity (Sullivan et al. 2011; Dai et al. 2007; Pantoliano et al. 1989).

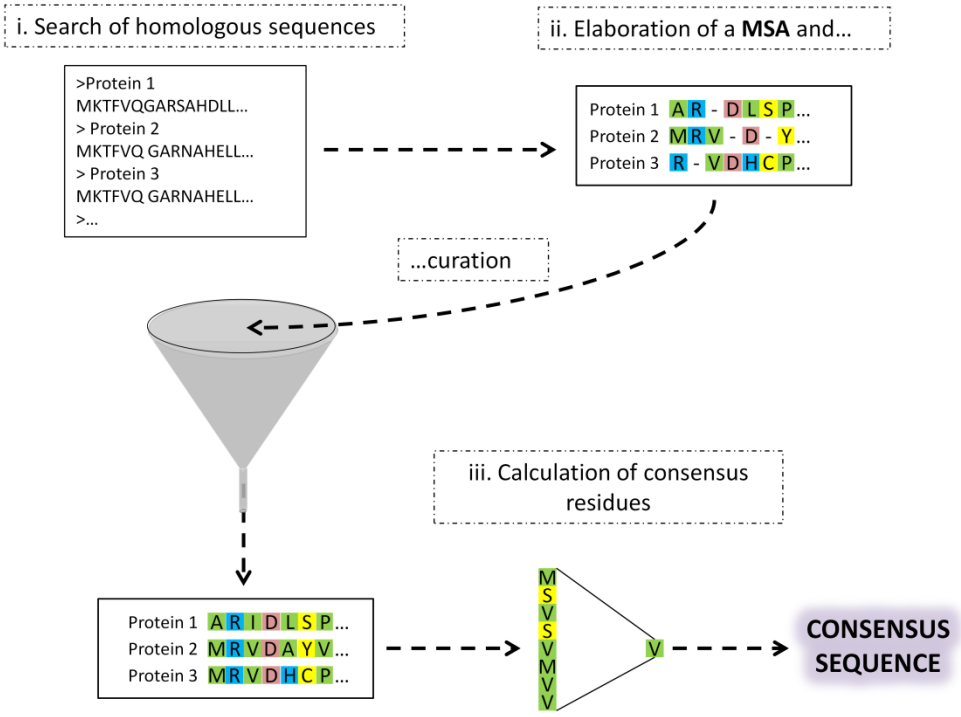


Figure 1.1.3 Schematic representation of the consensus method. Search of homologous sequences to the target protein, elaboration and curation of the MSA and calculation of the most repeated amino acid for each position to generate the consensus sequence.

It is well known that the main benefit of consensus design is the increment of stability, with differences in melting temperatures ranging from a few degrees up to 30ºC (Paatero et al. 2016; Dai et al. 2007; M. Lehmann et al. 2000; Wang et al. 2000; Wang et al. 1999; Wirtz and Steipe 1999; Nikolova et al. 1998). This is of special interest in directed evolution studies: given that the boundaries in protein evolvability are defined by the capacity to allocate beneficial -but destabilizing- mutations, the preservation of kinetic and thermodynamic stability (*i.e.* folding and function) is paramount to obtain successful designs. Thus, the insertion of consensus mutations can help to find shortcuts travelling the evolution route (Khersonsky et al. 2012; Bershtein and Tawfik 2008; Zahnd et al. 2007). Over the years, the combination of directed evolution and consensus design has proved to be a valuable protein engineering strategy (Saab-Rincón et al. 2018; Anbar et al. 2012; Komor et al. 2012;; Loening et al. 2006). For instance, directed evolution of Kemp eliminase towards the elimination of non-actividated benzisoxazoles was not possible due to the low stability of the enzyme, but the insertion of a set of consensus mutations allowed the enzyme to restart the evolution (Khersonsky et al. 2012). Another appealing benefit of using consensus approach is that it tends to break with the stablished trade-off

between activity and stability by rendering mutations usually located far from the catalytic core and not functionally involved (Porebski and Buckle 2016; Polizzi et al. 2006; Lehmann et al. 2002). Last but not least, the potential application of consensus design on immunogenicity is worth to be mentioned. This affair is still uncharted, but the possibilities in the field of proteins therapeutics deserve a second glance (De Groot et al. 2011; De Groot and Scott 2007; Blatt et al. 1996).

In the present Doctoral Thesis, we have applied a homemade consensus design to improve the stability of high-redox potential laccase evolved in the laboratory for functional expression in yeast (Mate et al. 2010).

1.1.2 Ancestral Sequence Reconstruction (ASR) and Protein Resurrection

The first ASR note was dated in 1963 when Pauling and Zuckerkandl stated that *"…it is possible to determine, with some probability, the amino-acid sequence of their presumed common polypetide-chain ancestor"/ When a fossil record is available, knowledge about the organisms concerned will go far beyond what has so far been believed possible*." (Pauling and Zuckerkandl 1963). Nowadays, the increment in the number of available sequences together with the progress in bioinformatics and molecular biology has made ASR a reality. Inferred ancestral sequences can be resurrected (*i.e.* functionally expressed within a modern host) and used to glance into the past for a better understanding of evolution while their potential implication in biotechnology cannot be underestimated. Indeed, resurrected enzymes usually display improved traits in terms of stability, promiscuity and evolvability, and as such, they could be suitable blueprints for laboratory evolution studies as reflected in the current Doctoral Thesis (Risso et al. 2018; Ayuso-Fernández et al. 2017; Babkova et al. 2017; Zakas et al. 2017; Whitfield et al. 2015; Cox and Gaucher 2014; Risso et al. 2013), **Figure 1.1.4**.

A reliable ASR experiment follows several steps:

i) Search of homologous extant sequences. The query of sequences may be carried out using the same databases as those used for consensus design. When the number of available sequences is overwhelming, a threshold of 30-90% identity can be applied in order to discard sequences with high similarity. The appropriate number of sequences relies on the mutation rates and time span of the reference protein. Consequently, some groups of sequences for ASR consist only a few tens while others may include hundreds (Merkl and Sterner 2016b; Perez-Jimenez et al. 2011; Yokoyama et al. 2008).

ii) Elaboration and curation of the MSA. The very same tools and procedures than those applied to consensus design are generally used whereas for areas containing uncertain alignments to be discarded, GBLOCKS can be applied (Castresana 2000).

iii) Deduction of a phylogenetic tree. This step is commonly achieved by using a Maximum Likelihood (ML) (Pupko et al. 2000; Yang et al. 1995) or a Bayesian method (Huelsenbeck and Bollback 2001). In ML the tree with the highest probability is selected, whereas with a Bayesian approach the preferred tree will show the highest posterior probability, (*i.e.* extracted from the prior expected probabilities of the trees). For ML approaches, Phylogenetic Analysis by Maximum Likelihood (PAML) (Yang 1997; Yang et al. 1995), Genetic Algorithm for Rapid Likelihood Inference (GARLI) (Bazinet et al. 2014) or Phylogenetic Analysis Using Parsimony (PAUP) (Swofford 1984) can be used. In order to apply the Bayesian methodology, MrBayes (Ronquist and Huelsenbeck 2003), PhyML (Guindon and Gascuel 2003) and PhyloBayes (Lartillot et al. 2009) are the programs of choice. Along these lines, literature can be also employed to infer the phylogenetic tree of a given protein (Yokoyama and Radlwimmer 2001).

iv) Reconstruction of the ancestral sequence: ML or Bayesian methodology is also employed in the final step. By using the selected extant sequences and the phylogenetic tree, the most probable sequences are inferred by ML. Conversely, the Bayesian methodology takes into account the most probable ancient sequence supported by the extant ones. In both cases, in addition to the sequences of extant proteins and a phylogenetic tree, a substitution model is needed. Substitution models are based on codons (Yang and Nielsen 2000; Yang et al. 2000) or on the amino acid sequence level, including Jones substitution model (JTT) (Jones et al. 1992) and WAG matrix (Whelan and Goldman 2001; D. T. Jones et al. 1992). These substitution models are performed to estimate the probability for a residue/nucleotide to be replaced by another one.

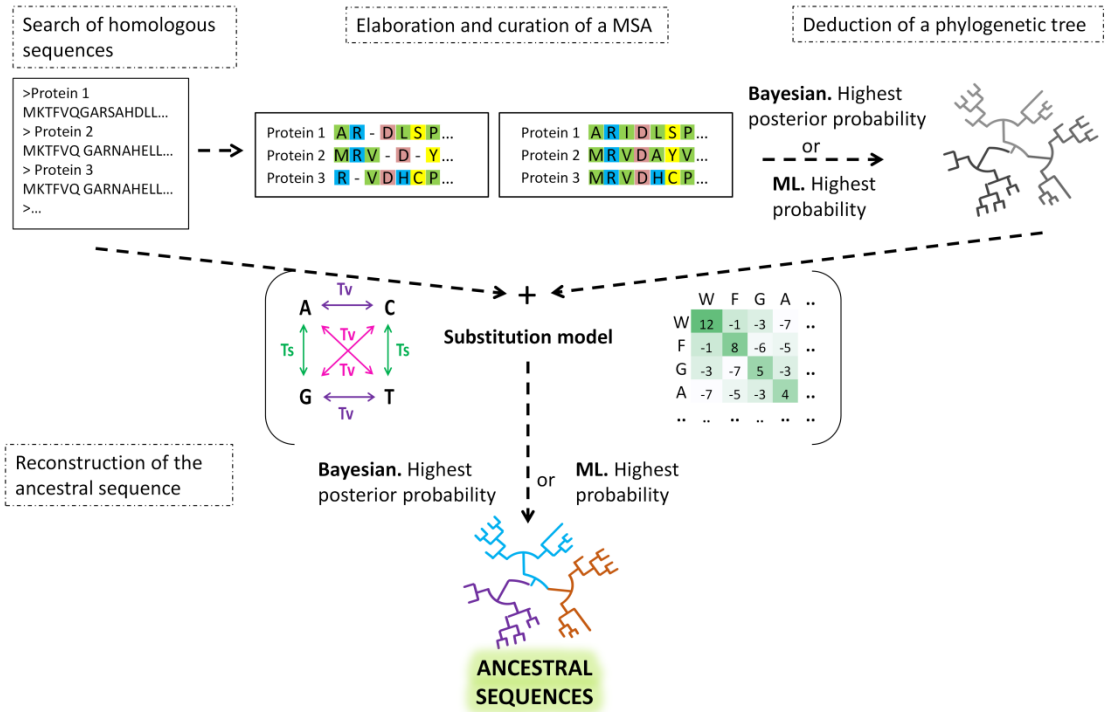


Figure 1.1.4 Schematic representation of the ancestral sequence reconstruction. The deduction of the phylogenetic tree is achieved by using the MSA and Bayesian/ML method. The combination of the sequences, the phylogenetic tree and a substitution model (amino acids or nucleotides) give rise to the reconstructed nodes sequences.

The information extracted after ASR can be harnessed in different forms. The most obvious is synthesis of the DNA sequences corresponding to the different ancestral nodes along the evolutionary pathway of a given protein family (Risso et al. 2013; Perez-Jimenez et al. 2011). An alternative to ASR is the design of a library with ancestral mutations for a modern enzyme by the Ancestral Mutation Method (AMM). With this approach, while preserving some desired properties of the extant protein, additional features provided by the ancestral mutations may arise (Yamashiro et al. 2010; Watanabe et al. 2006; Miyazaki et al. 2001).

As it has been above mentioned, ancestral enzymes usually display properties of biotechnological interest. High thermostability is the most common observed feature of ancestral enzymes. The explanation may stem from the fact that the temperature of the Precambrian earth (approximately 30ºC over current temperature) forced the adaptation of organisms for survival (Akanuma 2017; Garcia et al. 2017; Risso et al. 2014; Gaucher et al. 2008; Robert and Chaussidon 2006). There are striking examples in literature about Precambrian proteins with strong thermostabilities (Nguyen et al. 2017; Kratzer et al. 2014; Akanuma et al. 2013; Risso et al. 2013; Perez-Jimenez et al. 2011), reaching in some cases melting temperatures of 100ºC with differences of more than 30ºC respecting to their

modern counterparts. Along with the thermostabilization effect, the harsh conditions of the ancient earth may have driven some resurrected enzymes to be more adapted to work at extreme pH (Merkl and Sterner 2016b; Perez-Jimenez et al. 2011; Walker 1983). Another interesting trait of ancestral enzymes is promiscuity. Although current enzymes are specialized catalysts that play determined physiological roles in the cell (Copley 2015, 2017), many of them does seem to retain some latent/promiscuous activities, a reminiscent of their forgotten ancestral roles (Risso et al. 2018; Molina-Espeja et al. 2016). Indeed, ancient organisms couldn´t allow themselves to synthesis a full set of specific enzymes for every single cell task due the high metabolic cost that this would imply, yet they relied the majority of the functions on only a few generalist and promiscuous enzymes (Akanuma 2017; Alcalde 2017b; Risso et al. 2014; Akanuma et al. 2013; Jensen 1976). This is very well illustrated in the work of Risso and co-workers (Risso et al. 2013). They resurrected four Precambrian β-lactamases that showed active in the hydrolysis of a variety of second generation antibiotics not recognized by the modern counterpart. Another appealing feature of resurrected ancestral enzymes is increased expression (Manteca et al. 2017; Zakas et al. 2017a). Ancestral enzymes were very likely pressed to fold in the absence of chaperons, or in the best of the cases, in the presence of poorly efficient chaperons; hence, the ability to fold correctly mostly depended on the sequence protein itself (Risso et al. 2018; Candel et al. 2017).

We believe that the alliance of ASR with directed evolution may result appropriate to boost potential new features and/or to design finest models whereby allocating disruptive mutations and/or new activities (Risso et al. 2018; Alcalde 2017b; Merkl and Sterner 2016a; Cox and Gaucher 2014), **Figure 1.1.5**.

In this Doctoral Thesis ASR has been combined with directed evolution in two different enzymes, a laccase and a Rubisco from different origins, fungi and bacteria, respectively, as well as from different time lapse.

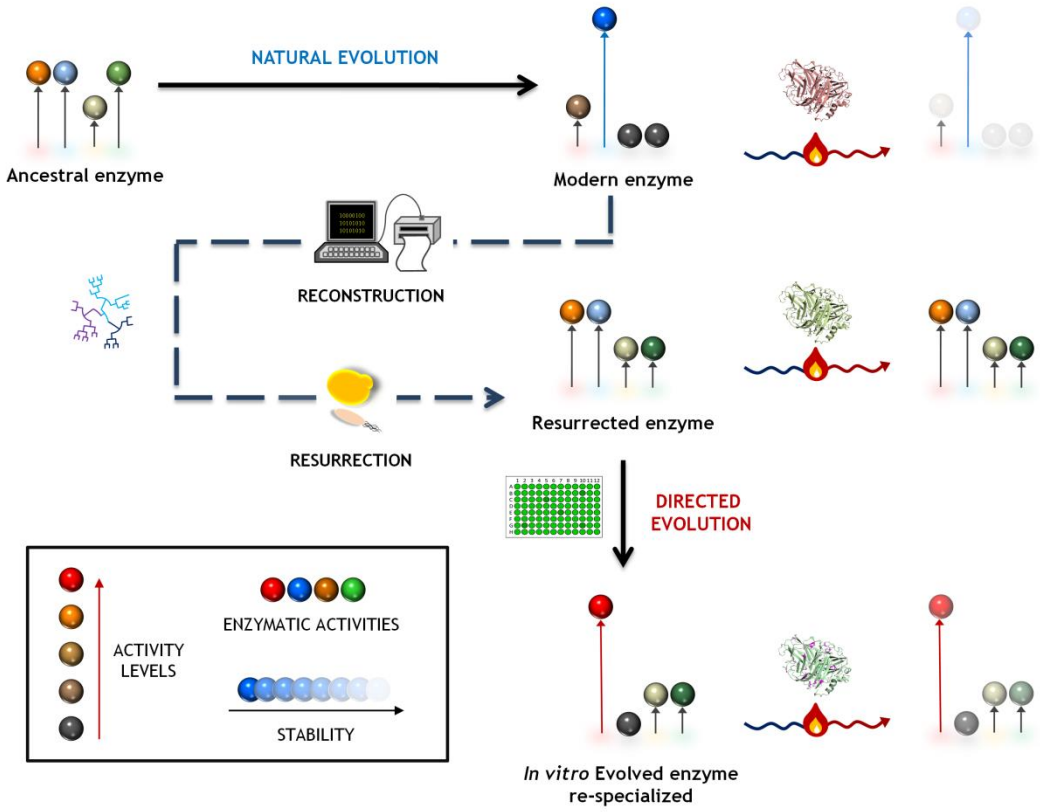


Figure 1.1.5 Directed evolution of ancestral enzymes. From a modern enzyme scaffold, a plausible approximation to the ancestral node is predicted. The sequence of the ancestral node is then synthesized and functional expressed (resurrected) in a suitable host (*e.g.* bacteria or yeast). Thereafter, the resurrected enzyme is subjected to directed evolution towards latent activities those are hardly shown by the modern counterpart. The strong stability of the resurrected node can help alleviate the detrimental effect of using high mutational loads during the directed evolution campaign. Figure adapted from (Alcalde 2017b).

1.2 Rubisco

1.2.1 General aspects

Ribulose-1,5-bisphosphate carboxylase/oxygenase, Rubisco (EC 4.1.1.39) is the most abundant enzyme in the biosphere, with a broad distribution in plants, algae and bacteria (Andersson and Backlund 2008; Ellis 1979). The main role of this enzyme of Precambric origin (~3500 millions of years ago) is the fixation of atmospheric CO_2 into biomass through the Calvin-Benson-Bassham reductive pentose phosphate cycle (Portis and Parry 2007; Benson and Calvin 1948). Accordingly, the main Rubisco activity is the carboxylation of ribulose-1,5-bisphosphate (RuBP) yielding two molecules of 3-phosphoglycerate (3-PGA). 3-PGA is necessary for the carbohydrate synthesis that is bounded to the growth of the organism. In general, Rubisco performs this reaction with turnover rates as low as 1 to 10 s^{-1}, affecting the photosynthesis and growth of organisms,

such as plants and algae. In order to overcome this hurdle, Rubisco-containing organisms usually produce large amounts of Rubisco (50% of the total soluble proteins in rice and wheat leaves) (Sharwood 2017; Robert E. Sharwood et al. 2016a; Carmo-Silva et al. 2015).

There are four forms of Rubisco sorted as I, II, III, and the *Rubisco-like Proteins*, (RLP) that is commonly considered as the form IV. RPL are not able to fix CO_2, although their participation in the enolization of RuBP analogues and thiosulphate oxidation has been described (Tabita et al. 2008; Portis and Parry 2007). Such classification is mainly driven by different combinations of large (L) and small (S) subunits. Form I Rubiscos are the most abundant group (present in high plants, algae and cyanobacteria) and show a hexameric holoenzyme composed of eight large subunit of ~50 kDa and eight small subunits of ~14-17 kDa (L_8S_8) (Andersson and Backlund 2008; Tabita et al. 2007). Rubiscos from chemoautotrophic bacteria, some dinoflagellates and photosynthetic proteobacteria constitute the form II, carrying only a dimer of the L subunit (L_2) (Mueller-Cajar et al. 2007; Tabita 1999) the minimum structure necessary for Rubisco activity. Form III Rubiscos can be found in Archaea and assembled as L_2 or $(L_2)_5$. Their main function is not related with CO_2 fixation but with metabolism of RuBP in archaean nucleotide salvaging pathway (Wilson et al. 2016; Sato et al. 2007) **Figure 1.2.1**

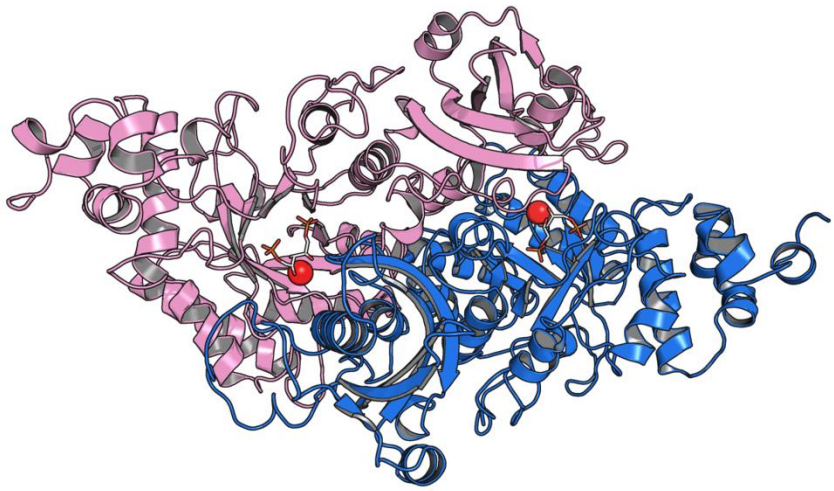


Figure 1.2.1 The Rubisco form II from *Rhodospirillum rubrum* (RubRr) L_2 used in the present Doctoral Thesis. The Rubisco structure is represented in cartoon with subunits in purple and blue. RuBP is shown in white and orange sticks and Mg^{2+} as red spheres (PDB code 9RUB).

Although this classification highlights the differences in Rubisco structures, the high degree of conservation in the catalytic mechanism and the conformation of the active site must be considered. Indeed, it has been stated that all Rubisco share a common

ancestor and this can be easily appreciated in the tertiary structure of the L subunit and the disposition of amino acids that belong to the catalytic region which are quite conserved among different Rubiscos (Andersson and Backlund 2008; Andersson and Taylor 2003; Tabita 1999), with a canonical structural arrangement in the dimer L_2. The C-terminal domain of one L-subunit and the N-terminal domain of the other L-subunit conform two catalytic sites (Andersson and Backlund 2008) **Figure 1.2.1**. By contrast, the S subunit is more divergent among Rubisco species and it seems not to be essential for activity (Andersson and Taylor 2003; Gutteridge 1991; Andrews 1988). In terms of reaction mechanism, all Rubiscos need to be activated prior catalysis. The activation starts via carbamylation of a conserved lysine in the catalytic pocket by binding CO_2 (Cleland et al. 1998; Taylor and Andersson 1996; Hartman and Harpel 1994; Schreuder et al. 1993). The resulting anionic carbamate is then stabilized by Mg^{2+}, generating a ternary complex of Enzyme-CO_2-Mg^{2+} (ECM) (Wilson and Whitney 2017; Jordan and Chollet 1983; McCurry et al. 1981).

Once that Rubisco is activated the catalysis of RuBP starts following several sequential events (Andersson 2008; Kellogg and Juliano 1997; Hartman and Harpel 1994).

i) After binding RuBP, an enediol-intermediate is produced through the desprotonation of the C3 of RuBP.

ii) This enediol provide a nucleophilic site at C2 that is attacked by a molecule of CO_2 (carboxylation) or O_2 (oxygenation).

iii) The resultant ketone suffers a hydration on the hydroxyl group of the C3.

iv) Finally, the C-C bond cleavage and the protonation process produce either two 3-PGA (carboxylation) or one 3-PGA and one 2-phosphoglycolate (2-PG) (oxygenation).

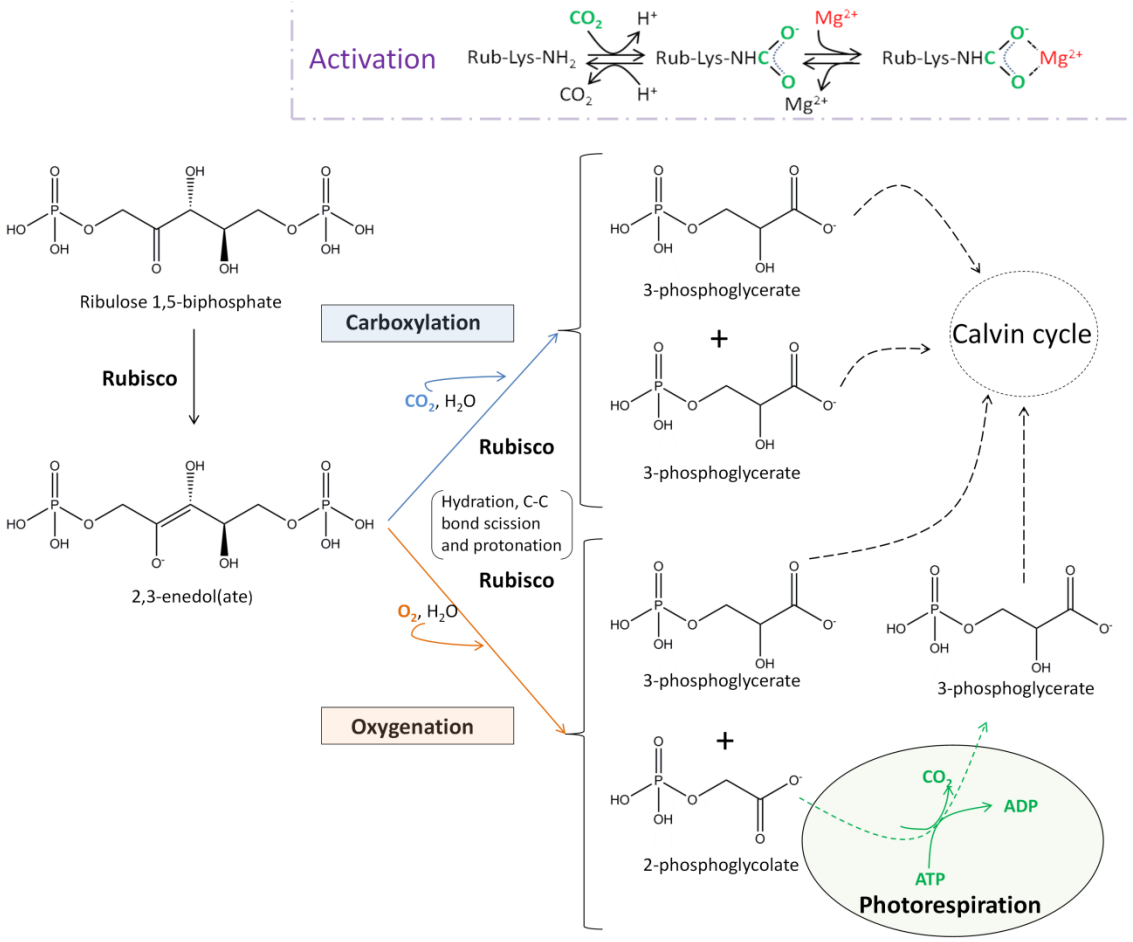


Figure 1.2.2 Schematic overview of rubisco activation and main reactions for the carboxylation and oxygenation of RuBP. The CO_2 and Mg^{2+} involved in the formation of the ternary complex ECM are highlighted in green and red, respectively. The routes for carboxylation and oxygenation are depicted in blue and orange, respectively. The regeneration of 2-PG through photorespiration is outlined as a green circle and the Calvin Cycle is represented as a dotted circle.

The catalytic bi-functionality inherent to Rubisco´s action mechanism implies that it cannot discriminate correctly between CO_2 and O_2 due to: i) their close electrostatic nature, ii) their unbalanced proportion in present atmosphere ($525O_2$:$1CO_2$) and iii) the direct binding of the co-substrate to the RuBP enediol intermediate without the formation of an enzyme Michaelis complex. As such, the tight competition between these two gases in Rubisco catalysis results in the undesirable side oxygenation reaction of RuBP to 3-PGA and 2-PG. The latter is a compound toxic to plant that needs to be recycled back to 3-PGA through a high energy demanding photorespiration, at the expense of adenosine triphosphate (ATP) and Nicotinamide adenine dinucleotide phosphate (NADPH) equivalents and the release of NH_3 and fixed CO_2 (for C3 crops it is assumed that, on average, ~25 % of the photosynthetic output is lost via Rubisco´s oxygenation) (Sharwood et al. 2016a; Whitney et al. 2011; Parry et al. 2003), **Figure 1.2.2**.

Research has been focused on the study of different Rubiscos that, despite having a strongly conserved catalytic mechanism, display different yields and capacity to differentiate between both gases (Sharwood et al. 2016b; Price and Howitt 2014; Andrews and Whitney 2003; Whitney and Andrews 1998).

1.2.2 Rubisco engineering

Given the Rubisco´s catalytic constrains, the last decade has witnessed many laboratories worldwide spending efforts and resources aimed at engineering more efficient Rubiscos in a drive towards a new, more productive agriculture (Long Stephen et al. 2015; Evans 2013; Parry et al. 2013;). Indeed, the most promising results have arrived by the hand of directed evolution. All these studies on directed Rubisco evolution are assisted by *in vivo* -genetic- screenings based on photosynthetic selection systems (*e.g. Rhodobacter capsulatus* and *Chlamydomonas reinhardtii*) and above all, on Rubisco dependent *E. coli* (RDE) selections like the MM1-*prk* RDE system (Antonovsky et al. 2017; Wilson and Whitney, 2017; Whitney et al. 2011; Mueller-Cajar and Whitney 2008a). With these strategies, forms I, II and III Rubisco have been evolved to give rise to an array of variants with mostly improved solubility and slightly modified kinetics (Wilson et al. 2016; Durão et al. 2015; Cai et al. 2014; Zhu et al. 2010; Satagopan et al. 2009; Mueller-Cajar and Whitney 2008b; Greene et al. 2007; Mueller-Cajar et al. 2007; Smith and Tabita 2003). During these evolution campaigns, outstanding metabolic engineering efforts have been done to wire Rubisco within the host metabolism, but limitation in terms of the reduced cell viability, the poor selection fidelity and the low throughput are still serious obstacles that remain to be circumvented (Antonovsky et al. 2017; Wilson and Whitney 2017; Wilson et al. 2016). More significantly, genetic selection methods can be only circumscribed to evolve the natural activity of Rubisco. As the enzyme is tightly linked to the cell survival, this approach precludes the improvement of biophysical properties like stability, a trait intrinsically connected to the protein evolvability in terms of mutational tolerance, fitness and biological function (Bloom et al. 2006).

In this Doctoral Thesis, we report a robust directed *in vitro* evolution platform that permits to evolve modern and ancestral Rubiscos unplugged from the host physiology.

1.3 Laccase

1.3.1 General aspects

Laccases (benzenediol:oxygen oxidoreductases, EC 1.10.3.2) belong to the group of multi-copper oxidases together with ceruloplasmin, ascorbate oxidase, bilirubin oxidases and ferroxidases (Baldrian 2006). Due to their low specificity, laccases are able to oxidize a wide variety of substrates using molecular oxygen from air and releasing water as only by-product (Cañas and Camarero 2010; Morozova et al. 2007a; Solomon et al. 1996). The natural role of laccases is directly connected to the producer organism, with a widespread distribution in plants (involved in the cross-linking of phenolic compounds on lignin and wounds (Mayer and Staples 2002), insects (where they participate in the cuticle sclerotization) (Claus 2004; Kramer et al. 2001), fungi (involved in lignin degradation and/or polymerization, morphogenesis and stress protection) (Brijwani et al. 2010; Alcalde 2007) and bacteria, (participating in pigmentation, morphogenesis, toxin oxidation and UV-resistance) (Singh et al. 2011; Martins et al. 2002; Hullo et al. 2001).

Laccases contain one Type 1 copper (T1Cu) or paramagnetic blue copper, involved in the binding and oxidation of reducing substrates and a trinuclear Cu cluster (T2/T3) for the reduction of oxygen to water (Alcalde 2007; Piontek et al. 2002; Solomon et al. 1996). These copper sites are organized within three β-barrel cupredoxin domains present not only laccases but in all multicopper oxidases, suggesting the existence of a common ancestor (Rydén and Hunt 1993). T1Cu is allocated in the domain 3 while the T2/T3 cluster is between domains 1 and 3. The three domains are preserved by two/three disulfide bridges (Kallio et al. 2011; Matera et al. 2008; Morozova et al. 2007b) **Figure 1.3.1**. The differences in the redox potential at the T1Cu site (E_{T1}) among laccases let these enzymes to be classified as low- medium- and high- redox potential laccases (from roughly 400 up to 790 mV *vs.* normal hydrogen electrode, NHE) (Shleev et al. 2005). Laccases from fungal basidiomycetes include the vast majority of high-redox potential lacasses (HRPL), whilst plant and bacterial laccases frequently display lower E_{T1}. At first glance, these differences in the redox potential are inconsistent with a similar structural arrangement around T1Cu, with many factors to be taken in consideration, such as the distance between T1Cu and the nitrogen atoms of coordinating histidines, the hydrophobicity in the area of T1Cu or the solvent accessibility, or the axial ligand to name a few (Osipov et al. 2014; Cambria et al. 2012; Hadt et al. 2012; Hong et al. 2011; Marshall et al. 2009; Tadesse et al. 2008; Durao et al. 2006; Li et al. 2004; Piontek et al. 2002; Xu et al. 1999).

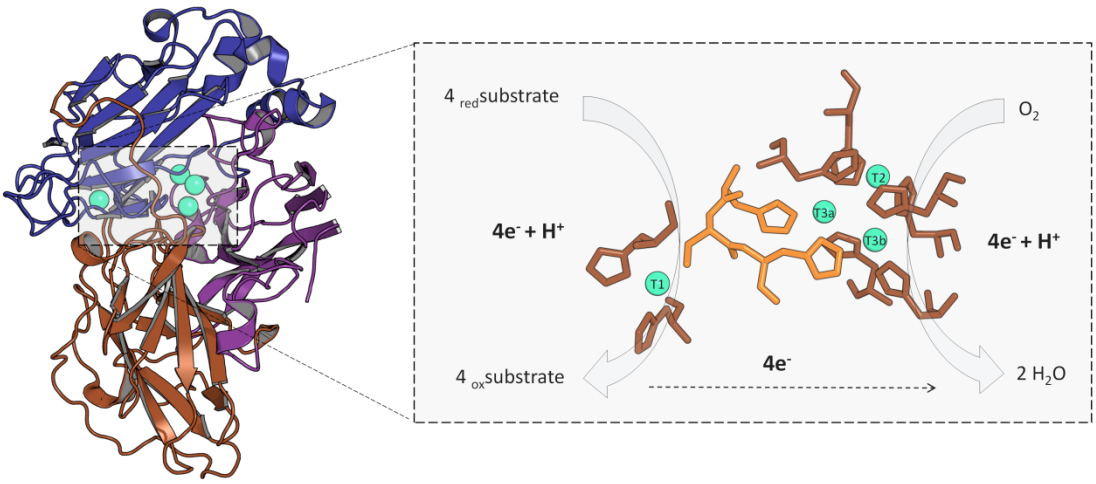


Figure 1.3.1 Molecular model of laccase from PM1L basidiomycete (OB-1 variant, used in the current Doctoral Thesis) with intramolecular electron transfer pathway. Domain 1 (D1), blue, Domain 2 (D2), brown. Domain 3 (D3), purple. Cooper atoms are represented as blue spheres. His and Cys residues are depicted as sticks. His-Cys-His tripeptide is highlighted in orange (PDB code 5ANH).

Although some mechanisms about laccase catalysis are still not fully elucidated, there are certain aspects already consolidated (Alcalde 2007; Morozova et al. 2007b).

i) The one-electron oxidation of the substrate takes place in the T1Cu that switches to its reduced state. The free cationic radical obtained as a result can undergo either polymerization or hydration.

ii) The T1Cu is rapidly re-oxidized by transferring the electron through the conserved tripeptide His-Cys-His. The electrons are accumulated by laccase that acts as a "battery".

iii) Once four substrate molecules have been oxidized, the four resultant electrons are used for the reduction of an oxygen molecule. This process takes place in the T2/T3 Cu cluster. After the formation of an intermediate that has been identified as a peroxide, two molecules of water are obtained, emerging by a different solvent channel than the oxygen entrance.

1.3.2 The laccase mediator system: Initiators

Laccases oxidize a broad range of substrates such as ortho- para- diphenols, meta substituted phenols, aromatic amines, Mn^{2+} (in presence of chelators) as well as various xenobiotic compounds (Rodgers et al. 2010; Morozova et al. 2007a). The variety of substrates available for laccases is highly dependent on the redox potential of the

compound. Despite this substrate range, high-redox potential compounds are hardly oxidized by laccase. This hurdle can be overcome by the use of redox mediators, small molecules that act as diffusible electron shuttles between the laccase and the target substrate through the laccase mediator system (LMS). Once it is oxidised by the enzyme, the mediator diffuses away from the enzymatic pocket and in turn oxidises any substrate that, due to its size, cannot directly enter the enzymatic pocket. Alternatively, the oxidized mediator could rely on an oxidation mechanism not available to the enzyme, thereby extending the range of substrates accessible, including high-redox potential compounds. By definition, an ideal mediator should be a laccase substrate whose oxidized and reduced forms are stable but must not inhibit the enzymatic reaction. In addition, its redox conversion should be cyclic. Ideally, without side reactions, a redox mediator can perform many cycles without degradation (Hilgers et al. 2018; Canas and Camarero 2010; Morozova et al. 2007a; Galli and Gentili 2004; Alcalde et al. 2002), **Figure 1.3.2A**. This general concept is true for many redox mediators that can follow different routes (*e.g.* the electron transfer route (ET) for ABTS; the hydrogen-atom transfer (HAT) route, for 1-hydroxybenzotriazole (HBT), or violuric acid (VLA) and the ionic route for (2,2,6,6-Tetramethylpiperidin-1-yl)oxyl (TEMPO), **Figure 1.3.3**. However, there are other relevant mediators (known as initiators and studied in the present Doctoral Thesis), which follow different principles and that are involved in the polymerization of vinyl monomers (*e.g.* styrene, acrylamide) (Hollmann and Arends 2012; Singh et al. 2000).

Classically, the polymerization of vinyl monomers is a complex task that can be sparked by azo compounds or peroxides, yet the process is not very efficient and leaves potentially hazardous residues in the final product (Rodriguez et al. 2014; Singh et al. 2000; Priddy 1994; Mark et al. 1963). As an environmentally friendly alternative, the enzymatic-initiation of radical polymerization harnesses the LMS comprised by an initiator, typically a β-diketone and the laccase: once the initiator is oxidized interacts with the monomer starting the polymerization process but unlike classical mediators, initiators are included in the final polymer structure, which precludes its reuse (Hollmann and Arends 2012; Durand et al. 2000; Lizotte and Long 2003; Teixeira et al. 1999; Derango et al. 1992), **Figure 1.3.2B**. Several studies in which LMS is employed for polymerization of vinyl monomers have been described, but with limited efficiency due to the poor activity of laccase towards initiators (Yang et al. 2015; Hollmann and Arends 2012; Hollmann et al. 2008; Tsujimoto et al. 2001; Ikeda et al. 1998). In this Doctoral Thesis we have tackled the directed evolution

of an ancestral laccase to accomplish the oxidation of the β-diketone 1,3-cyclopentanedione, a poor substrate for its modern counterpart.

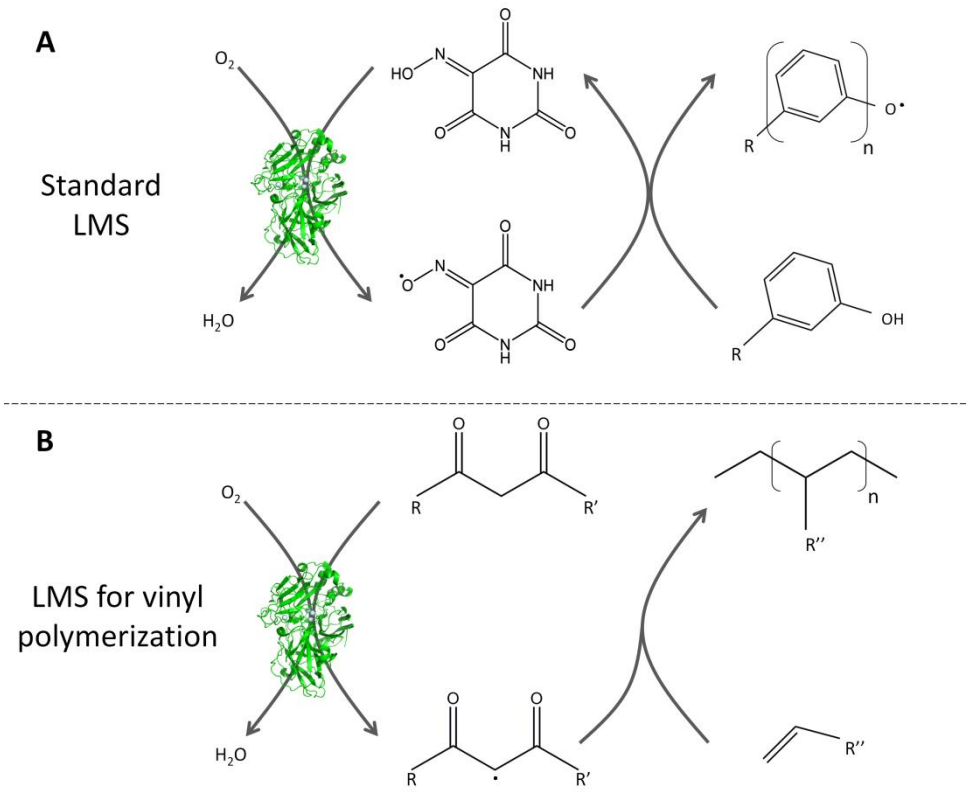


Figure 1.3.2 Schematic representation of laccase mediator system. A) Standar LMS where mediator is regenerated after oxidized the substrate. **B)** LMS when the mediator is a β-diketone that once oxidized is able to start the polymerization of vinyl monomers getting also included in the polymer structure.

1.3.3 Laccase applications and engineering

Laccases are known as ideal green catalysts because of their minimal requirements and wide substrate range (Canas and Camarero 2010; Rodgers et al. 2010; Morozova et al. 2007a; Riva 2006; Solomon et al. 1996). As such, it is not surprising to find commercial laccases for different industrial purposes such as in food, textiles, pulp and paper, pharma, biofuels or cosmetics (Shin et al. 2019; Cannatelli and Ragauskas 2017; Mate and Alcalde 2017; Viswanath et al. 2014; Virk et al. 2012; Osma et al. 2010; Mikolasch and Schauer 2009). From the treatment of corks surface aimed at reducing the extraction of agents into the wine to their implementation into biosensors for analytic uses, laccase applications are spreading more and more within modern society, (Naveen et al. 2019; Mate and Alcalde 2017; Di Fusco et al. 2010; Milligan and Ghindilis 2002; Cliffe et al. 1994). But many of these case stories have been achieved thank to the adaptation of laccases to the harsh industry conditions.

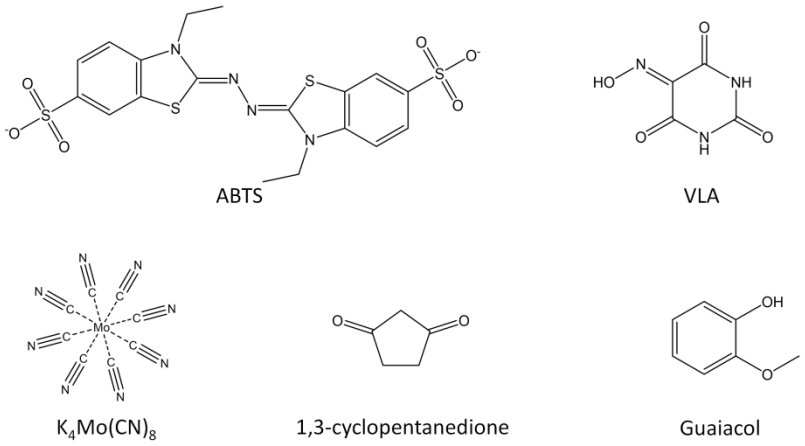


Figure 1.3.3 Structures of redox mediators employed in this Thesis.

Indeed, protein engineering, and particularly directed evolution (Mate and Alcalde 2015) have been documented to improve activity (Pardo et al. 2016; Santiago et al. 2016; Toscano et al. 2013), pH stability (Sheng et al. 2017; Maté 2013; Torres-Salas et al. 2013), heterologous expression (Camarero et al. 2012; Mate et al. 2010; Koschorreck et al. 2009; Bulter et al. 2003), stability at high temperatures (Isabel Pardo et al. 2018; Scheiblbrandner et al. 2017; Mollania et al. 2011) in organic solvents (Rasekh et al. 2014; Zumárraga et al. 2007), in ionic liquids (Liu et al. 2013) or even to human blood (Maté 2013). More recently, directed laccase evolution has been combined with computational methods in order to boost the E_{T1} (Mateljak et al. 2019a) as well as to design a family of thermostable chimeras (Mateljak et al. 2019b).

In the present Doctoral Thesis the study of laccase was two-fold: first, a consensus method was developed and applied to insert consensus mutations to enhanced stability without jeopardize activity and second, ancestral laccase resurrection and directed evolution was applied to improve activity against initiators.

2 OBJECTIVES

Despite directed evolution and ancestral resurrection are extremely potent protein engineering methods, they have never been combined for generating improved biocatalysts. In this Doctoral Thesis, we have brought together both approaches aimed at opening new venues for protein engineers, using as enzyme templates two very different enzymes, Rubisco and laccase.

Accordingly, the main objectives of this Doctoral Thesis were:

1. The reconstruction, resurrection and evolution of Precambrian Rubiscos from the extant form II Rubisco from *Rhodospirillum rubrum*. The resurrected Rubisco and the modern counterpart were evolved *in vitro* to enhance thermostability and their offspring was carefully characterized.
2. To resurrect Paleozoic nodes of fungal laccases as well as to set out the platform for their directed evolution towards initiators.
3. To design fungal laccases through consensus mutagenesis. By adding ancestral reconstruction to this pool, consensus/ancestor mutations proved to be valuable for their functional expression, activity and stability.

3 MATERIALS AND METHODS

3.1 Reagents and materials

Table 3.1.1 Reagents Chapter 1

Reagents	Company
2-phosphoglyceric acid (2PG)	Acros Organics
3-phosphoglyceric acid (3PG)	Sigma
Adenosine triphosphate (ATP)	Sigma
Agarose	BioRad
Dye binding Pierce Coomassie Plus	Thermo Fisher
Gelcode Blue stain	Thermo Fisher
Imidazole	Sigma
Isopropyl B-D-thiogalactopyranoside (IPTG)	Thermo Fisher
Magnesium chloride	Sigma
Manganese (II) chloride	Sigma
N-amylamine	Acros Organics
Nicotinamide adenine dinucleotide hydride (NADH)	Sigma
NuPAGE Bis-Tris Gels	Thermo Fisher
Protease inhibitor mixture for bacterial cells	Sigma
Ribulose 1,5-bisphosphate (RuBP)	Santa Cruz Biotechnology
Sodium bicarbonate	Sigma
Sodium chloride	Sigma
Tris	Sigma
UltimaGold Scintillant	Packard Bioscience
β-Mercaptoethanol	Sigma

Table 3.1.2 Reagents Chapter 2 and 3

Reagents	Company
1,3-cyclopentanedione	Sigma
2, 6-dimethoxyphenol (DMP)	Sigma
2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS)	Sigma
Agarose	BioRad
Bis-Tris	Fluka
Calcium suplhate dihydrate	Sigma
Copper (II) sulfate pentahydrate	Sigma
Ethanol	Panreac
Guaiacol	Sigma
Magnesium chloride	Sigma
Phosphoric acid	Sharlau
Potassium hydroxide	Panreac
Potassium octacyanomolybdate	Sigma
Potassium phosphate monobasic	Sigma
Potassium sulfate	Sigma
Sinapic acid	Sigma
Sodium chloride	Sigma
Violuric acid	Sigma
β-Mercaptoethanol	Sigma

Table 3.1.3 Medium components.

Medium components	Company
Amino acids supplements without uracil	Difco (BD)
Ampicilin (Amp)	Sigma
Bacto Agar	BD
Bacto peptone	BD
Chloramphenicol	Sigma
D-(+)-Galactose	Sigma
D-(+)-Glucose	Sigma
D-(+)-Rafinose pentahydrate	Sigma
Uracil	Sigma
Yeast extract	BD
Yeast nitrogen base without amino acids (YNB)	Difco (BD)

Table 3.1.4 Molecular biology kits.

Molecular biology kits	Company
Deoxyribonucleotide triphosphate (dNTP)	Sigma
Gel Loading Solution	Sigma
Gel Red DNA stain	BioRad
GeneRuler 1Kb DNA Ladder	Thermo Scientific
GeneRuler 100pb Plus DNA Ladder	Thermo Scientific
Gibson ASssembly™ Master Mix	New England Biolabs
NucleoSpin plasmid® Kit	Macherey Nagel
Protein assay dye reagent Kit II (Bradford)	BioRad
Yeast Transformation Kit	Sigma
Zymoclean Gel DNA recovery Kit	Zymoresearch
Zymoprep Yeast Plasmid Miniprep Kit	Zymoresearch

Table 3.1.5 Strains and plasmids.

Strains and plasmids	Company
Bacterial expression vector pTrcHis2 B	Thermo Fisher
E. coli XL1-Blue competent cells	Stratagene
E. coli XL2-Blue competent cells	Stratagene
Expression shuttle vector pJRoC30	Caltech
S. cerevisiae strain BJ5465	LGC Promochem

Table 3.1.6 Commercial enzymes.

Commercial enzymes	Company
Bovine serum albumin (BSA)	Sigma
DNase I (RNase-free)	New England Biolabs
Glyceraldehyde 3-phosphate dehydrogenase (GAPDH)	Sigma
Glycerol-3-phosphate dehydrogenase (GPDH)	Sigma
Mutazyme II	Agilent
Pfu-ultra High Fidelity DNA Polymerase	Agilent
Taq DNA Polymerase	Sigma
Triosephosphate isomerase (TPI)	Sigma
BamHI restriction enzyme	New England Biolabs
NcoI-HF restriction enzyme	New England Biolabs
SalI restriction enzyme	New England Biolabs
Phosphoglycerate kinase (PGK)	Sigma
XhoI restriction enzyme	New England Biolabs

3.2 Culture media

Composition of all mediums described here is referred to a final volume of one liter of distilled water. All mediums and solutions, if not specified otherwise, were autoclaved 30 min at 121°C.

3.2.1 Culture media for bacteria (*E. coli*) growth

Table 3.2.1 Luria-Bertani (LB) medium with ampicillin (LB/Amp).

Bacto Peptone	10 g
Yeast Extract	5 g
NaCl	10 g
Ampicilin (100 mg/mL)	1 mL

LB/Amp medium for the selective growth of *E. coli* cells transformed with pJRoC30 and pTrcHis2 B vectors which contain ampicillin resistance gene (Sambrook, Fritsch, and Maniatis 1989). First, extracts and salt were dissolved in distilled water and pH was adjusted to 7.0. After the sterilization by autoclaving it was necessary to wait until temperature decrease to 50°C and thereafter, filtered ampicillin was added. In order to prepare solid medium, 20 g of agar were added before sterilization.

Table 3.2.2 Luria-Bertani medium with ampicillin (LB/Amp) and IPTG.

Bacto Peptone	10 g
Yeast Extract	5 g
NaCl	10 g
IPTG (100 mM)	3 mL
Ampicilin (100 mg/mL)	1 mL

LB/Amp with IPTG medium for the selective growth of *E. coli* cells transformed with and pTrcHis2 B vector which contains ampicillin resistance gene and an IPTG inducible promoter for protein expression (Sambrook, Fritsch, and Maniatis 1989). First, extracts and salt were dissolved in distilled water and pH was adjusted to 7.0. After the sterilization by autoclaving it was necessary to wait until temperature decrease to 50°C and thereafter, filtered ampicillin and IPTG were added.

Table 3.2.3 Luria-Bertani medium with ampicillin (LB/Amp) and IPTG in flask format.

Bacto Peptone	10 g
Yeast Extract	5 g
NaCl	10 g
IPTG (100 mM)	10 mL
Ampicilin (100 mg/mL)	1 mL

LB/Amp with IPTG medium for the selective growth of *E. coli* cells transformed with pTrcHis2 B vector which contains ampicillin resistance gene and an IPTG inducible promoter for protein expression (Sambrook, Fritsch, and Maniatis 1989). First, extracts and salt were dissolved in distilled water and pH was adjusted to 7.0. After the sterilization by autoclaving it was necessary to wait until temperature decrease to 50°C and thereafter, filtered ampicillin and IPTG were added.

Table 3.2.4 Super optimal broth medium (SOB).

Bacto Peptone	2 g
Yeast Extract	0,5 g
NaCl	0,05 g
KCl	1 mL

Extracts and salts were dissolved in distilled water, pH was adjusted to 7.0 and sterilized by autoclave.

Table 3.2.5 Super optimal broth medium with catabolite repression (SOC).

SOB medium	5 mL
$MgCl_2$ (2 M)	25 µL
Glucose (20% p/v)	100 µL

SOC medium is used for *E. coli* transformation (Sambrook, Fritsch, and Maniatis 1989). $MgCl_2$ and glucose were previously filter-sterilized and thereafter mixed with SOB medium. For each transformation it was necessary to prepare a new SOC solution.

3.2.2 Culture medium for yeast (*S. cerevisiae*) growth

Table 3.2.6 YP medium (1.55 x).

Bacto Peptone	30,77 g
Yeast Extract	15,38 g

YP medium is used in laccase expression medium. The components were dissolved in water and the medium was sterilized by autoclave.

Table 3.2.7 YPD medium.

Bacto Peptone	20 g
Yeast Extract	20 g
Glucose (20% p/v)	100 mL
Chloramphenicol (25 mg/mL)	1 mL

YPD medium for yeast growth. After the components were dissolved in water, the medium was sterilized by autoclave. After sterilization it was necessary to wait until temperature decrease to 50°C and thereafter, filtered glucose and chloramphenicol were added. In order to prepare solid medium, 20 g of agar were added before sterilization.

Table 3.2.8 Minimal liquid medium.

YNB medium (67 g/L)	100 mL
Amino acids Supplements (10 x)	100 mL
Rafinose (20% p/v)	100 mL
Chloramphenicol (25 mg/mL)	1 mL

Selective medium without uracil for the growth of *S. cerevisiae* cells transformed with pJRoC30 (+ gene), which contains gene ura3 that complements for uracil auxotrophy. Water is previously autoclaved and after the temperature decrease to 50°C the remaining components -previously sterilized by filtration were added.

Table 3.2.9 Minimal solid medium (SC drop-out plates).

Bacto agar	20 g
YNB médium (67 g/L)	100 mL
Amino acids Supplements (10 x)	100 mL
Glucose (20% p/v)	100 mL
Chloramphenicol (25 mg/mL)	1 mL

Selective medium without uracil for the growth of *S. cerevisiae* cells transformed with pJRoC30 (+ gene), which contains gene ura3 that complements for uracil auxotrophy. It is prepared as described above for the minimal liquid medium, but including agar in the autoclave sterilization.

Table 3.2.10 Laccase expression medium in flask format.

YP medium (1,55 X)	720 mL
Buffer KH2PO4 (1 M, pH 6)	67 mL
CuSO4 (1 M)	2mL
Ethanol (100%)	30,9 mL
Galactose (20% p/v)	110 mL
Chloramphenicol (25 mg/mL)	1 mL

KH_2PO_4 buffer, $CuSO_4$, galactose and chloramphenicol were sterilized by filtration.

Table 3.2.11 Selective expression medium for laccases (SEM).

YNB medium (67 g/L)	100 mL
Amino acids supplements (10 x)	100 mL
Galactose (20% p/v)	100 mL
Buffer KH_2PO_4 (1 M, pH 6)	67 mL
Ethanol 100%	31 mL
$CuSO_4$ (1 M)	1 mL
Chloramphenicol (25 mg/mL)	1 mL

SEM was used for laccase expression in microplate format (96-well) during HTS. KH_2PO_4 buffer, $CuSO_4$, galactose, YNB, amino acids supplements and chloramphenicol were sterilized by filtration.

3.3 Methodology employed in Chapter I: Directed *-in vitro-* evolution of Precambrian and extant Rubiscos.

3.3.1 ASR and resurrection

Alignment, phylogeny and ancestral sequence reconstruction

Bacterial Rubisco homologues were retrieved from the December 2013 release of the complete genomes database available at the NCBI (www.ncbi.nlm.nih.gov/genome). This search produced 295 protein sequences. We discarded sequences that were too long or too sort, resulting in a set of 201 sequences. The sequences were aligned using MUSCLE (available at https://www.ebi.ac.uk/Tools/msa/muscle/) and a distance matrix was then generated using one minus the sequence similarity as the parameter to assess the evolutionary distance between two sequences. The distribution for the calculated distances revealed three different groups and we found that several taxa were represented in the three groups of sequences, specifically: cyanobacteria; firmicutes; and α, β and γ proteobacteria. The sequences belonging to the group closest to the query were retained for sequence reconstruction resulting in 136 sequences. Since Mr. Bayes efficiency depends strongly on sequence set size, we selected 46 sequences for phylogenetic analysis. We made sure, however, that this subset provided a uniform coverage of the cyanobacteria, firmicutes, α-, β - δ and γ-proteobacteria taxa. A literature search confirmed that all the selected sequences belong to Rubisco type II. An Archean sequence was used as outgroup to generate a rooted tree making a total of 47 sequences, **Table 3.3.1**. The tree topology and the branch lengths of the trees were estimated from these sequences using the Bayesian method implemented in version 3.1.2 of the program MRBAYES (available at http://mrbayes.sourceforge.net/). This analysis used the Jones substitution model and two independent Markov-chain Monte Carlo runs, each with four chains, and it was performed for 3,603,600 generations to ensure adequate convergence (0.037). The nodes obtained in the tree had posterior probabilities higher than 0.8. However, we targeted the three oldest nodes of the tree for sequence reconstruction, which had a probability of essential unity, **Figure 3.3.1**.

Table 3.3.1 Rubisco sequences used for ancestral reconstruction. GI numbers were accessed from GenBank.

GI number	
332295273	198283921
260893699	344199398
295697414	121604879
428217181	171060437
428211462	257092457
428299658	543964646
427713058	78484769
428221690	336476213
22299049	148244796
261406831	118602694
337749006	316936114
386724715	39937699
317154242	115524692
436842937	90424392
89900184	86748076
296136113	91975534
410693894	294677366
74318656	126464741
291612494	77465277
261855695	146280021
340782403	384263572
390952417	83593735
71909249	83311795
431932595	

The sequence reconstruction was performed using PAML version 4.4e. (available at http://abacus.gene.ucl.ac.uk/software/paml.html) and with the WAG evolutionary model (available at https://www.ebi.ac.uk/goldman-srv/WAG/), **Figure 3.3.2**. Subsequently, the proteins encoded by the most probable sequences at each of the three nodes targeted were prepared in the laboratory (*i.e.*, "resurrected") and characterized experimentally.

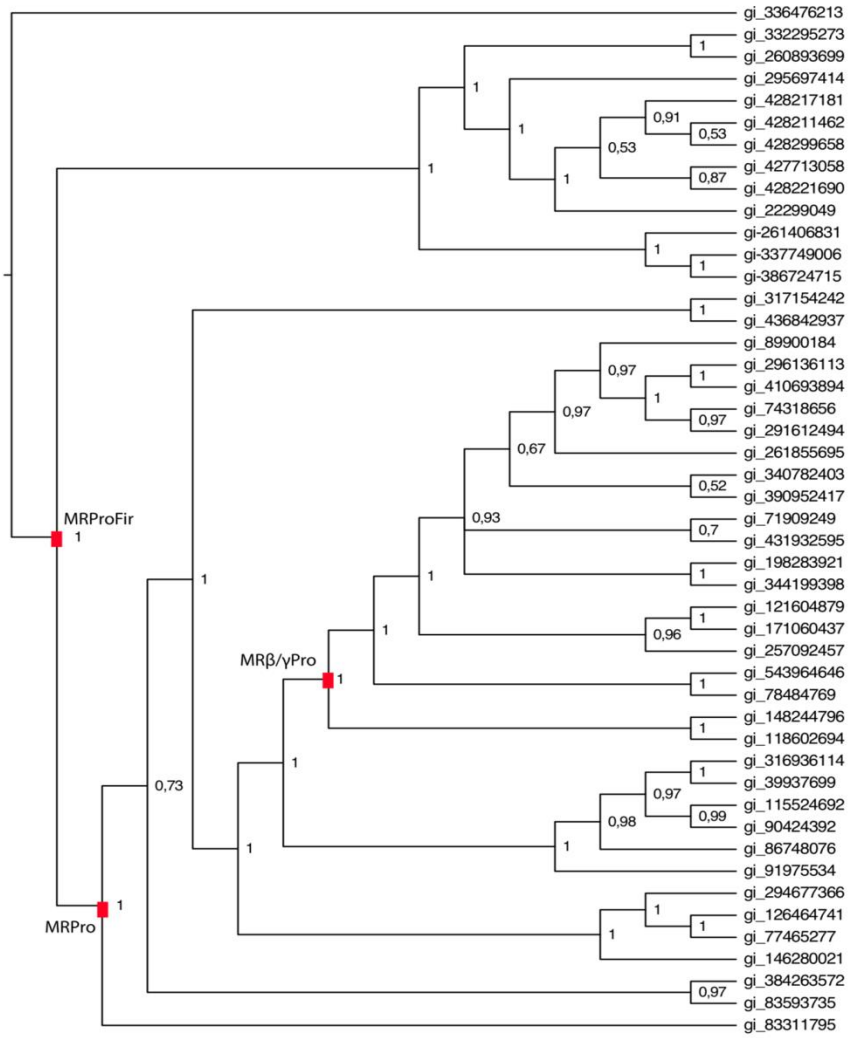


Figure 3.3.1 Inferenced evolutionary trees with the posterior probability at each node.

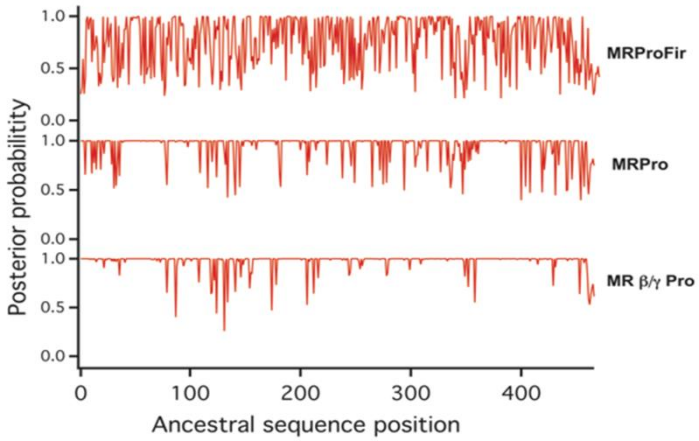


Figure 3.3.2 Robustness of the reconstructed ancestors as measure of the distribution of posterior probability.

Genes and synthesized reagents

RubRr (GeneBank: X00286.1), MRProFir, MRPro and MRβ/γPro were synthesized by ATG biosynthetics (Merzhausen, Germany). RuBP unlabeled and [1-^{3}H]RuBP-labeled, were synthesized and purified as described (Kane et al. 1998). [^{14}C]2-Carboxyarabinitol-1,5-biphosphate ([2-^{14}C]CABP) was synthesized as described (Pierce, et al. 1980).

Protein modelling

The structural model of wild-type RubRr at a resolution of 2.6 Å (1 Å=0.1 nm) (PDB code 9RUB) was used to map the mutations and for comparison with its ancestral counterparts. MRProFir, MRPro and MRβ/γPro were modelled by the Phyre2 server (Protein Homology/analogY Recognition Engine V 2.0) (Kelley and Sternberg 2009) available at www.sbg.bio.ic.ac.uk/phyre2.

3.3.2 Directed evolution

Library creation methods

General aspects. For each generation PCR fragments were cleaned, concentrated and loaded onto a low melting point preparative agarose gel (Bio-Rad, Hercules, CA), and then purified using the Zymoclean gel DNA recovery kit (Zymo Research, Orange, CA). PCR products were cloned into the plasmid pTrcHis2 B from Invitrogen (California, USA) fused to a 6xHis-tag at the C-terminus using the Gibson Assembly Master Mix. pTrcHis2 B was linearized at the corresponding restriction sites with NcoI-HF and Sal I, then cleaned and purified as described for PCR products.

Error prone PCR (*ep*PCR) for adaptive evolution, focused evolution and neutral genetic drift was carried out using a gradient thermocycler (Mycycler, Biorad, USA). The cycling parameters unless others specified were: 94ºC for 2 min (1 cycle), 94ºC for 0.45 min, 50ºC for 0.30 min, 72ºC for 1.30 min (28 cycles) and 72ºC for 10 min (1 cycle). The cycling parameters for high fidelity PCR in focused evolution were: 94ºC for 2 min (1 cycle), 94ºC for 0.30 min, 55ºC for 0.30 min, 72ºC for 1.30 min (28 cycles) and 72ºC for 10 min (1 cycle). The primers used for amplification unless others specified were RubDir sense and RubRev antisense for RubRr; 62Dir sense and 62Rev antisense for MRPro, **Table 3.3.2**.

Table 3.3.2 List of primers used in error prone PCR (*ep*PCR)

PCR RubRr	
RubDir	5' TAAATAAGGAGGAATAAACCATGGACC 3'
RubRev	5' TGATGATGATGATGATGGTCCGCCG 3'
PCR MRPro	
62Dir	5' CGATTAAATAAGGAGGAATAAACCATGGATCAGTCTAAACGTTACG 3´
62Rev	5' CTCAATGATGATGATGATGATGGGCAGGCAGGGCGCTTGCC 3'

*ep*PCR with Taq/$MnCl_2$: To find an appropriate mutational rate, several mutant libraries were prepared at different concentrations of $MnCl_2$ in a 50 µL final volume containing: pTrcHis2 B-RubRr 2,82 ng/µL, 90 nM RubDir, 90 nM RubRev, 0.3 mM Deoxyribonucleotides (dNTPs) (0.075 mM each), 3% (v/v) dimethylsulfoxide (DMSO), 1.5 mM $MgCl_2$, increasing concentrations of $MnCl_2$ (0.01, 0.02, 0.03 and 0.05 mM) and 0.05 U/µL Taq polymerase. For adaptive evolution, focused evolution and genetic drift, independent mutant libraries of RubRr and MRPro were prepared in a 50 µL final volume with pTrcHis2 B-RubRr 0.055 ng/µL or pTrcHis2 B-MRPro 0.055ng/µL, 90 nM of each primer, 0.3 mM dNTPs (0.075 mM each), 3% (v/v) dimethylsulfoxide (DMSO), 1.5 mM $MgCl_2$, 0.03 mM $MnCl_2$ and 0.05 U/µL Taq polymerase.

*ep*PCR with GeneMorphII: To find an appropriate mutational rate, several mutant libraries were prepared with different concentrations of pTrcHis2 B-RubRr (50, 100, 200 and 750 ng), 37 nM RubDir, 37 nM RubRev, 0.8 mM dNTPs (0.2 mM each), 3% (v/v) dimethylsulfoxide (DMSO), and 0.05 U/µL Mutazyme II.

Focused evolution by Mutagenic Organized Recombination Process by Homologous *In vivo* Grouping (MORPHING): The MORPHING method (36) was adapted to *E. coli* with several modifications. Four different PCRs were performed for each library, two high fidelity PCRs upstream and downstream the mutagenic block, one mutagenic PCR (loop 6 or catalytic pocket) and one assembly PCR. The primers used for the upstream high fidelity PCR were RubDir sense and RubCatRev2 or RubL6Rev2 antisense for MORPHING in the catalytic pocket or the loop 6 of RubRr, respectively; 62dir sense and 62CatRev2 or 62L6Rev2 antisense for MORPHING in the catalytic pocket or the loop 6 of MRPro, respectively. The primers used for the downstream high fidelity PCR were RubRev antisense and RubCatDir2 or RubL6Dir2 sense for MORPHING in the catalytic pocket or the loop 6 of RubRr, respectively; 62rev antisense and 62CatDir2 or 62L6Dir2 sense for

MORPHING in the catalytic pocket or the loop 6 of MRPro, respectively, **Table 3.3.3**. Upstream and downstream high fidelity PCR reactions contained 0.25 μM of each primer, 1 mM dNTPs (0.25 mM each), 3% (v/v) dimethylsulfoxide (DMSO), 0.2 ng/μL of pTrcHis2 B-RubRr or pTrcHis2 B-MRPro template and 0.05 U/μL Pfu polymerase. The primers used for the mutagenic PCR were: RubCatDir1 or RubL6Dir1 sense and RubCatRev1 or RubL6Rev1 antisense for MORPHING in the catalytic pocket or the loop 6 of RubRr, respectively; 62CatDir1 or 62L6Dir1 sense and 62CatRev1 or 62L6Rev1 antisense for MORPHING in the catalytic pocket or the loop 6 of MRPro, respectively. Mutagenic PCR reactions contained 0.09 μM of each primer, 0.3 mM dNTPs (0.075 mM each), 0.37 ng/μL of pTrcHis2 B-RubRr or pTrcHis2 B-MRPro template, 3% (v/v) dimethylsulfoxide (DMSO), 1.5 mM $MgCl_2$, 0.1 mM $MnCl_2$ and 0.05 U/μL Taq polymerase. The assembly PCR was done in two PCR-steps. The cycling parameters for the first part of the assembly PCR were 98ºC for 2 min (1 cycle), 98ºC for 0.30 min, 45ºC for 0.30 min, and 72ºC for 2 min (15 cycles). This PCR reaction contained 0.25 μM of each primer, 1 mM dNTPs (0.25 mM each), 1ng/μL of each template (PCRs products from the two high fidelity PCRs and the mutagenic one), 3% (v/v) dimethylsulfoxide (DMSO), and 0.05 U/μL Pfu polymerase. After the last cycle all components were added again in the PCR tube except the templates and then the second part of the assembly PCR was performed as follows: 98ºC for 2 min (1 cycle), 98ºC for 0.30 min, 50ºC for 0.30 min, 72ºC for 2 min (28 cycles) and 72ºC for 10 min (1 cycle). This PCR reaction contained 0.25 μM of each primer, 1 mM dNTPs (0.25 mM each), 3% (v/v) dimethylsulfoxide (DMSO), and 0.05 U/μL Pfu polymerase. The primers used in both PCRs were RubDir sense and RubRev antisense for RubRr; 62Dir sense and 62Rev antisense for MRPro.

Genetic drift: Seven consecutive rounds of neutral genetic drift were performed on RubRr by *ep*PCR with Taq polymerase/$MnCl_2$, **Table 3.3.2**. Once the mutant libraries were prepared according to the protocol described in the next section, aliquots of 20 μL of clones with relative activity higher than 0.7 *vs.* parental type (measured with the HTS assay described below) were transferred with the help of a liquid handler station (Freedom EVO 100, TECAN, Männedorf, Schweiz) to new 96 deep well plates containing 100 μL of LB. After 13 h of incubation at 37ºC and 250 rpm and 80% relative humidity (humidity shaker Minitron-INFORS, Biogen, Spain), cultures were pooled together and plasmids were extracted with NucleoSpin Plasmid kit. This mix was then used as template to create a new generation of neutral drift. After 7 generations, 20 neutral clones were randomly selected and characterized.

Table 3.3.3 List of primers used for MORPHING.

UP HF PCR Cat/loop RubRr	
RubDir	5' TAAATAAGGAGGAATAAACCATGGACC 3'
RubCatRev2	5' GCACTTTCCACAGGGCCGAGATATTG 3'
Or	
RubL6Rev2	5' GGTATAGCCGCGCTTGGACTGG 3'
DW HF PCR Cat/loop RubRr	
RubRev	5' TGATGATGATGATGATGGTCCGCCG 3'
RubCatDir2	5' CTTGCGCGACACCATCGCCCT 3'
Or	
RubL6Dir2	5' AGGCCCAGGGGCCGTTCTACC 3'
UP HF PCR Cat/loop MRPro	
62Dir	5' CGATTAAATAAGGAGGAATAAACCATGGATCAGTCTAAACGTTACG 3´
62CatRev2	5' CTTTCCACATATGGGCAATATTGC 3'
Or	
62L6Rev2	5' AGAGTAGCCGCGCTTCGACTGG 3'
DW HF PCR Cat/loop MRPro	
62Rev	5' CTCAATGATGATGATGATGATGGGCAGGCAGGGCGCTTGCC 3'
62CatDir2	5' GACCATCCGATTAGTCGCTGATGCG 3'
Or	
62L6Dir2	5' GATGAGGCACAGGGCCCCTAT 3'
MUT PCR Cat RubRr	
RubCatDir1	5' CAATATCTCGGCCCTGTGGAAAGTGC 3'
RubCatRev1	5' AGGGCGATGGTGTCGCGCAAG 3'
MUT PCR Loop RubRr	
RubL6Dir1	5' CCAGTCCAAGCGCGGCTATACC 3'
RubL6Rev1	5' GGTAGAACGGCCCCTGGGCCT 3'
MUT PCR Cat MRPro	
62CatDir1	5' GCAATATTGCCCATATGTGGAAAG 3'
62CatRev1	5' CGCATCAGCGACTAATCGGATGGTC 3'
MUT PCR Loop MRPro	
62L6Dir1	5' CCAGTCGAAGCGCGGCTACTCT 3'
62L6Rev1	5' ATAGGGGCCCTGTGCCTCATC 3'

Preparing Mutant libraries

PCRs products were cleaned and concentrated. They were cloned behind the trc (trp-lac) promoter of the expression vector pTrcHis2 B using the Gibson Assembly Master Mix. *E.coli* XL1-Blue were transformed with the plasmid containing the RubRr, the MRPro or the corresponding mutant libraries, plated on LB-Amp agar and grown for 16 h at 37ºC. Individual colonies were picked into 44 mm deep well plates containing LB-Amp medium (0.1 mL). Each plate contained an internal standard with Rubisco parental type (column 7, rows A to H) and a negative control (lysate from *E.coli* transformed with pTrcHis2 B without Rubisco, column 6, rows A to D).

The plates (master plates) were incubated at 37ºC with shaking at 250 rpm and 80% relative humidity in a humidity shaker (Minitron-INFORS, Biogen, Spain). After 16 h, clones from this pre-culture were inoculated (using a cryo-replicator CR1000 from Enzyscreen, Haarlem, Netherlands) into deep well plates with fresh LB-Amp (lysate plates). After 5 hours of incubation at 37ºC, 250 rpm and 80% relative humidity, LB-Amp with IPTG was added (0.4 mL final volume). The plates were incubated at 25ºC with shaking at 250 rpm and 80% relative humidity for 16 h. The plates were centrifuged at 3000 *g*, the medium was discarded and the cell pellets were frozen at -80ºC. After 12 h the frozen cell pellets were re-suspended in a lysis mixture of Tricine-KOH buffer (0.4 mL, 100 mM, pH 8.0) containing lysozyme (0.25mg/mL), DNase I (1.5 U/mL) and 10 mM $MgCl_2$. After 60 min at 37ºC and 250 rpm the lysates were centrifuged at 3000 *g* and the supernatant was used for the dual HTS assay as well as for the thermostability screening protocol.

Dual HTS assay

60 µL of the supernatant from lysate plates were transferred with the help of a liquid handler station (Freedom EVO 100, TECAN, Männedorf, Schweiz) to the reaction plates. 50 µL of activation buffer (20 mM $MgCl_2$ and 40 mM $NaHCO_3$ in 100 mM Tris-KOH buffer pH 8.0) was added to each well of the reaction plate. Reaction plates were briefly stirred and after 15 min at room temperature 70 µL of the Rubisco reaction buffer were added. The reaction buffer contained 20 mM $MgCl_2$, 40 mM $NaHCO_3$, 0.5 mM RuBP (for spectrophotometric assay) or 3 mM RuBP (for High performance liquid chromatography- Evaporative light scattering detector (HPLC-ELSD) assay) in 100 mM Tris-KOH buffer, pH 8.0.

Spectrophotometric assay: The reaction plate was stirred briefly and after 4 min at room temperature 40 µL of each well were transferred to a new plate containing 40 µL of 80%

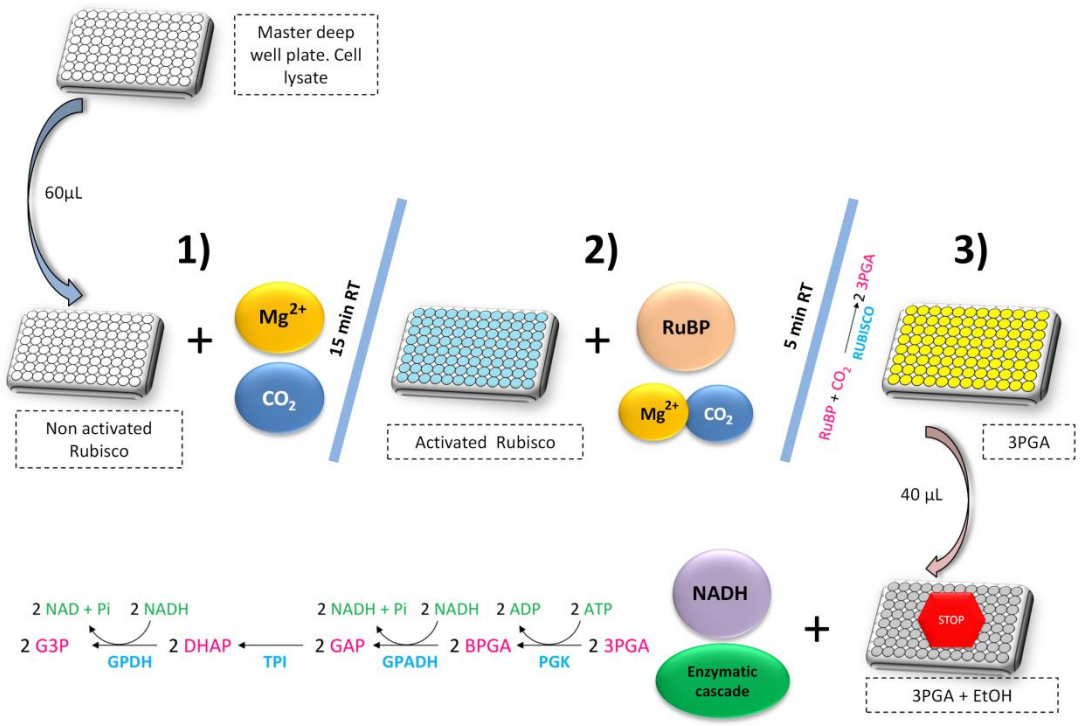


Figure 3.3.3 Depletion assay method. 1) Transfer and enzyme activation. The activation of Rubisco takes place upon the addition of the activation buffer. 2) The activated Rubisco is mixed with the reaction buffer and the carboxylation of RuBP starts. 3) 40 µL of the reaction plate is transferred to a new plate containing EtOH to stop the reaction. Thereafter, the NADH depletion buffer is added and the reaction is followed. Selected clones of the spectrophotometric assay are then transferred from the master plate to a new reaction plate and analysed by the HPLC-ELSD. Two consecutive re-screenings are performed (see further details in this section).

(v/v) ethanol to quench the reaction. The plates were stirred and after 5 minutes 120 µL of the NADH depletion buffer was added. Spectrophotometric-depletion buffer composition: 1.5 mM $MgCl_2$, 0.4 mM NADH, 0.5 mM ATP, 0.5 U/mL TPI, 0.5 U/mL αGPDH, 0.5 U/mL GAPDH and 5 U/mL PGK in 100 mM Tris-KOH buffer (pH 8.0). The depletion assay was followed in kinetic mode (340 nm) using a plate reader, **Figure 3.3.3** (SpectraMax Plus 384, Molecular Devices, Sunnyvale, CA).

HPLC-ELSD assay: Selected clones of the spectrophotometric assay were transferred from the lysate plate to a new reaction plate and analysed by the HPLC-ELSD (high-performance liquid chromatography with evaporative light scattering detector) assay. The reaction plate was stirred briefly and after 30 min at room temperature 400 mM acetic acid was added to stop the reaction. The HPLC conditions were 50% Buffer A (12 mM ammonium acetate, pH 6.0), 40% Buffer B (200 mM n-amylamine in buffer A, pH 6.0) and 10% C (acetonitrile). The final concentration of n-amylamine was 80 mM. The HPLC analysis was performed with a pump (Varian 9012, Canada) coupled to a Nucleosil C18

(250 x 4,6 mm, Sugelabor, Spain) kept at 30ºC. Autosampler (Hitachi L-2200, VWR, Spain) was set to 10ºC to prevent evaporation of the samples. The ELSD conditions were 118ºC and 3.2 L/min of nitrogen flow. Quantification was made with the aid of calibration curves of RuBP and 3PG. Chromatograms were analyzed with Varian Star v.6.41 software.

First re-screening: Aliquots of 10 μL of the best clones were removed from master plates to inoculate in 90 μL of LB-Amp in new 96 deep well plates. Columns 1 and 12 (rows A and H) were not used. After 16 h of incubation at 37ºC and 250 rpm and 80% relative humidity, 10 μL were transferred to the adjacent wells, containing 90 μL of fresh LB-Amp and further incubated for 5 hours in the same conditions. Then 300 μL of LB-Amp with IPTG were added. The plates were incubated at 25ºC with shaking at 250 rpm and 80% relative humidity for 16 h. Accordingly, every single mutant was grown in 4 wells. Parent types were subjected to the same procedure (lane D, wells 7-11). Plates were assessed using the same protocol of the screening described above.

Second re-screening: An aliquot from the best clones of the first re-screening was inoculated in 10 mL of LB-Amp and incubated at 37ºC and 250 rpm for 16 h. Plasmids from these cultures were extracted with NucleoSpin Plasmid kit and they were transformed together with the parental type into *E.coli*. After 16 h at 37ºC, five colonies of every single mutant were picked, inoculated in 100 μL of LB-Amp and re-screened as described above.

Screening assay for thermostability

RubRr and MRPro mutant libraries from adaptive evolution, focused evolution and a random sample of the outcome from the genetic drift campaign were subjected to a thermostability screening protocol. Lysate plates were duplicated with the help of the liquid handler station by transferring 85 μL of lysates to a thermocycler plate (Multiply PCR plate without skirt, neutral, Sarstedt, Germany) and 60 μL to the initial activity plate. Thermocycler plates were sealed with thermo resistant film (Deltalab, Spain) and incubated at 64ºC in the thermocycler (MyCycler, Bio-Rad Laboratories). Incubation took place for 10 min (so that the assessed activity was reduced ½ of the initial activity). Then, thermocycler plates were placed on ice for 10 min and further incubated for 5 min at room temperature. 60 μL of lysates were transferred from thermocycler plate to a new plate (residual activity plate) with the help of the robot. Each well of both residual and initial activity plates was filled with 50 μL of activation buffer (20 mM $MgCl_2$ and 40 mM $NaHCO_3$ in 100 mM Tris-KOH buffer pH 8.0) Plates were briefly stirred and after 15 min at room temperature 70 μL of the Rubisco reaction buffer was added and the plates

subjected to the spectrophotometric assay as described above. Relative activities were calculated from the difference between the absorption after incubation and that of the initial measurements normalized against the parental type in the corresponding plate. Thermostability values came from the ratio between residual activities (RA) and initial activities (IA) values. Selected clones were subjected to HPLC-ELSD analysis and to two consecutive re-screenings as described above.

DNA sequencing

Plasmid-containing mutant Rubisco genes were sequenced by GATC-Biotech. The samples were prepared with 5 µL of 100 ng/µL plasmid and 5 µL of 5 µM of each primer, Rub2 dir sense, Rub4 rev antisense, Rub5 dir sense and Rub6 dir sense for RubRr; 62 2 dir sense, 62 4 rev antisense, 62 5 dir sense and 62 6 dir sense for MRPro, **Table 3.3.4**.

Table 3.3.4 List of primers used for sequencing Rubiscos.

Seq RubRr	
Rub2 dir	5' TAACTTCTTGCATTATCACC 3'
Rub4 rev	5' CGAAGGGCTGATTGCCCTGG 3'
Rub5 dir	5' AAATTAAAGAGGTATATATTAATGTATCG 3'
Rub6 dir	5' CGCCAAGATGCACGATTTCTAT 3'
Seq MRPro	
62 2 dir	5' GCATGACTTCTACGTTCCTCCG 3'
62 4 rev	5' TTAATGAAATCGCCGCCGAGCCA 3'
62 5 dir	5' CAAAAGCTGGAGCTTGCAAGCTT 3'
62 6 dir	5' ATCGAGCCGGTCATGGTGCA 3'

3.3.3 Biochemical characterization

Production and purification

RubRr, MRPro and selected mutants were transformed into XL1-Blue and grown in 0.35 L flasks with LB-Amp at 37ºC, and at an optical density at 600 nm (OD_{600}) ~ 0.6. Rubisco production was induced with flask expression medium for 8 h at 30ºC before harvesting by centrifugation (5 min at 6,000 x *g*). The cells were then frozen for 12 h at -80ºC and suspended in the lysis mixture that was described above (supplied with 1 mg/mL of lysozyme and 2 U/mL of DNase I). Additionally, they were submitted to an ultrasonic lysis by sonication and harvested by centrifugation (1 h at 12,000 x *g*). The supernatant was dialyzed in binding buffer, Tris-HCl (100 mM, pH 7.8) containing 0.5 M NaCl and 0.01 M imidazole and then passed through a Ni^{2+} Sepharose colum (HisTrap FF 5mL, GE

healthcare, Sweden) attached to a fast protein liquid chromatography system (FPLC, Äkta Purifier, GE Healthcare Uppsala, Sweden). Bound protein was eluted with a gradient of elution buffer in Tris-HCl (100 mM, pH 7.8) containing 0.5 M NaCl and 0.3 M imidazole. Fractions were immediately dialyzed with successive changes to stability buffer (20mM Tris-HCl, pH 8.0, containing 50 mM NaCl) and stored at 4ºC.

Thermostability

The melting temperature (T_m) of purified Rubisco samples was measured in a VP-Capillary Differential scanning calorimetric (DSC) (Microcal, GE Healthcare, Northhampon, USA) following a protocol described elsewhere (Pierce, et al. 1980). Protein solutions for calorimetric experiments were exhaustively dialyzed against Hepes buffer (25 mM, pH 8.0) containing 50 mM NaCl and the buffer from the last dialysis step was used as reference in the calorimetric experiments. Prior to scans with protein solutions, several buffer-buffer baselines were obtained to ensure proper equilibration of the calorimeter. For each protein, experiments at a scan rate of 90 K/h and a protein concentration of 10 μM were typically performed. Concentrations were determined spectrophotometrically at 280 nm using extinction coefficients and molecular weights calculated from the corresponding sequences. No significant scanning rate effects on thermal denaturation were detected when other scan rates (60 and 180 K/h) or other protein concentrations (30 μM and 0.6 μM) were employed.

Analysis of expression

XL1Blue cells containing parental type RubRr, MRPro, and mutants were harvested after growth and induced as indicated above. The cells were suspended in extraction buffer (100 mM EPPS-NaOH, pH 8.0) containing 1 mM Ethylenediaminetetraacetic acid (EDTA), 20 mM $MgCl_2$, 2 mM dithiothreitol, 0.043% (w/v) protease inhibitor mixture for bacterial cells and lysed with a French pressure cell (140 MPa). An aliquot of lysate sample was mixed with an equal volume of SDS buffer (125 mM Tris-HCl, pH 6.8, 4% (w/v) SDS, 20% (v/v) glycerol, 150 mM 2-mercaptoethanol, 0.01% (w/v) bromophenol blue), and the remaining was centrifuged at 38,000 x g for 15 min at 4°C. Aliquots of the supernatant were either assayed for protein content using the dye binding Pierce Coomassie Plus kit, treated with an equal volume of SDS buffer (soluble protein sample for Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis), or incubated with 25 mM $NaHCO_3$ at 25°C for 30 min before measuring Rubisco content ([2-^{14}C]carboxyarabinitol-P2 (^{14}C-CABP) binding, see below) and carboxylase activity under substrate RuBP-limiting

and -saturating conditions (see below). For the specificity measurements the lysis was done as described above but without adding EDTA in the extraction buffer.

PAGE analysis. Proteins were separated by SDS-PAGE using 4–12% NuPAGE Bis-Tris Gels. The Bis-Tris gels were buffered with MES and electrophoresed at 200 V according to the supplier's instructions. Protein bands were visualized using Gelcode Blue reagent and unstained protein standard as the ladder.

Quantification of Rubisco content. Rubisco content in the XL1-Blue extracts was quantified by ^{14}C-CABP inhibitor binding after preincubation with 40 mM $NaHCO_3$ and 20 mM $MgCl_2$ as described (Pierce, et al. 1980). The binding assays were done in duplicate and incubated for 20 min with 15 or 45 μM ^{14}C-CABP at 25°C and the amount of Rubisco-bound ^{14}C-CABP was recovered by gel filtration and then mixed with one volume of scintillant for scintillation counting.

CO_2 kinetics

Rubisco affinity for CO_2 under no O_2 atmosphere at 25ºC (K_C) and catalytic constant for rates of CO_2-fixation (K_{cat}^C)was measured by $^{14}CO_2$ fixation at 25°C, pH 8.0, according to (Andrews 1988) in nitrogen sparged septum-capped scintillation vials. The assays were initiated by adding cell free soluble *E. coli* protein extracts (that had been preincubated for 20 min in buffer containing 20 mM $MgCl_2$ and 25 mM $NaHCO_3$) into N_2-equilibrated assay buffer EPPS-NaOH (100 mM, pH 8.0) containing 20 mM $MgCl_2$, 0.5 mM RuBP, 0.1 mg/mL carbonic anhydrase, and varying concentrations of $NaH^{14}CO_3$ (2.4-19 mM). The assays were stopped after 2 min with 0.2 volumes of 20% (v/v) formic acid and dried at 90°C, and then the residue was dissolved in 0.5 mL of water before adding 1 volume of scintillant for scintillation counting. The data was fitted to the Michaelis-Menten equation and kinetic parameters were obtained by dividing the extrapolated maximal carboxylase activity by the concentration of Rubisco active sites quantified from the ^{14}C-CABP analysis as described above, **Figure 3.3.4**.

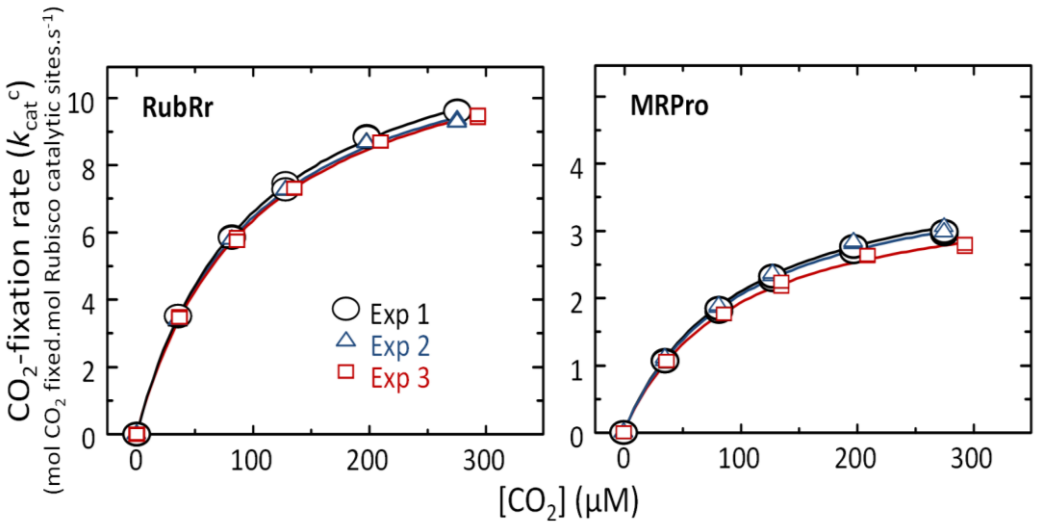


Figure 3.3.4 Measuring the CO_2-fixation rate and K_m for CO_2. Representative plots comparing the variation in the response of $^{14}CO_2$-fixation rate to differing [CO_2] for RubRr and MRPro under anoxic conditions at 25°C. The data from three independent experiments for each enzyme were fitted to the Michaelis-Menten equation (indicated by the lines shown) to derive the apparent K_m for CO_2 (K_C) and the maximal rate of carboxylation (V_C^{max}). The Rubisco-catalytic site content in each experiment was quantified by [^{14}C]-2-CABP use to normalise the extrapolated maximal carboxylase activities (k_{cat}^C) measured in each experiment.

RuBP kinetics

Cell free soluble *E. coli* protein extracts were preincubated with 25 mM $NaHCO_3$ for 20 min at 25 °C to activate the Rubisco. Carboxylase activities were measured using $NaH^{14}CO_3$ assays (Andrews 1988) containing different amounts of substrate RuBP (0–0.6 mM). Assays were buffered with EPPS-NaOH (100 mM, pH 8.0) containing 20 mM $MgCl_2$ and were performed in duplicate with unbuffered RuBP added to initiate catalysis. The assays were stopped after 70 seconds with 0.2 mL 20% (v/v) formic acid and transfer to heater block set at 90°C and dried. Then, the dried residue was dissolved in 0.5 mL of water before adding 1 volume of scintillant for scintillation counting. The K_m for RuBP was derived from fitting the data to the Michaelis-Menten equation.

CO_2/O_2 specificity

$S_{C/O}$ was measured according to the method of Kane and collaborators (Kane, et al. 1994) using purified Rubisco. Each assay (1 mL total volume) contained Rubisco and 20 µg carbonic anhydrase buffered in 30 mM triethanolamine, 15 mM Magnesium Acetate (pH 8.1). The assays were equilibrated for 60 min with a continuous flow of humidified gas containing the desired O_2/CO_2 mole-fraction ratio produced by mixing pure CO_2 with O_2 using a series of three Wösthoff precision gas-mixing pumps. The assays were initiated by

the addition of [C^{2}-[^{3}H]RuBP and after 30 minutes alkaline phosphatase added and the [^{3}H]-glycerate and [^{3}H]-glycolate recovered by anion exchange chromatography and separated by HPLC and $S_{C/O}$ calculated from the ratio of both products as described (Kane, et al. 1994).

3.4 Methodology employed in Chapter II: Ancestral Resurrection and Directed Evolution of Fungal Paleozoic Laccases

3.4.1 ASR and resurrection

Alignment, phylogeny and ancestral sequence reconstruction

Fungal laccases homologues were retrieved from the April 2015 release of the complete genomes database available at National Center for Biotechnology Information (NCBI) (Coordinators 2018). We deleted those sequences that did not include the four sequence regions identified by Kumar and coworkers (Kumar et al. 2003). These regions correspond to laccases signatures that differentiate them from the broader class of multi-copper oxidases. All sequences including terms such as "hypothetical", "predicted" or "putative" were discarded, as well of those sequences coming from insects or ascomycetes, resulting in a set of 120. The sequences were aligned using MUSCLE (available at https://www.ebi.ac.uk/Tools/msa/muscle/) and a distance matrix was then generated using one minus the sequence similarity as the parameter to assess the evolutionary distance between two sequences. The distribution for the calculated distances revealed two different groups and we found that the major clade of Basidiomycota, the Agaricomycotina, class Agaricomycetes was overrepresented over the rest. For phylogenetic analysis we employed the 87 sequences closest to the query since MrBayes efficiency relays strongly on sequence set size. Zygomycota and ascomycota sequences were also included as outgroups to generate a rooted tree making a total of 89 sequences, **Table 3.4.1**. The tree topology and the branch lengths of the trees were estimated from these sequences using the Bayesian method implemented in version MrBayes 3.1.2 (available at http://mrbayes.sourceforge.net/). This analysis used the JTT model and two independent Markov-chain Monte Carlo runs, each with four chains, and it was performed for 2113000 generations to ensure adequate convergence (0.04). The nodes obtained in the tree had a posterior probabilities higher than 0.8, with very few exceptions. We targeted the nodes 95, 98 and 100 in order to explore different alternatives of the tree, the three of them with probabilities close to the unity, **Figure 3.4.1**. The sequence reconstruction was performed using PAML version 4.4e (available at http://abacus.gene.ucl.ac.uk./software/paml.html) and with the WAG evolutionary model (available at http://www.ebi.ac.uk/goldman-srv/WAG/), **Figure 3.4.2**. Subsequently, the

proteins encoded by the most probable sequences at each of the three nodes targeted were synthesized to initiate the experiments for resurrected them (*i.e.* functionally express them).

Table 3.4.1 Laccase sequences used for ancestral reconstruction. GI numbers were accessed from GenBank.

GI number			
28268547	409151767	751715398	621090605
380704397	630209062	751711089	55670399
10801036	758348005	751681404	342161416
238632213	56785434	751672765	255523026
761951808	38194441	385282687	751000472
597970335	749841057	597933438	302672386
598003325	327349048	597933466	3273348
761926573	751056491	749880787	595781984
761926579	113207314	749885673	595781910
89274031	752354829	108936945	270047922
527300376	750965362	88687733	630355769
527300379	50724580	385141759	630350115
558633451	646305153	385139612	170101418
558633467	124495024	572695304	628856226
751689593	9957143	380750098	170101420
751689606	242220107	449546190	628856225
749897724	242210489	751007004	595781910
749897755	568441732	83415007	270047924
763724579	599119398	2833190	630359269
763725325	599096545	2833189	630350115
270485111	16041065	636611527	
348609404	34922426	46578391	
348609402	471887945	553303989	
409151775	471872355	553303985	

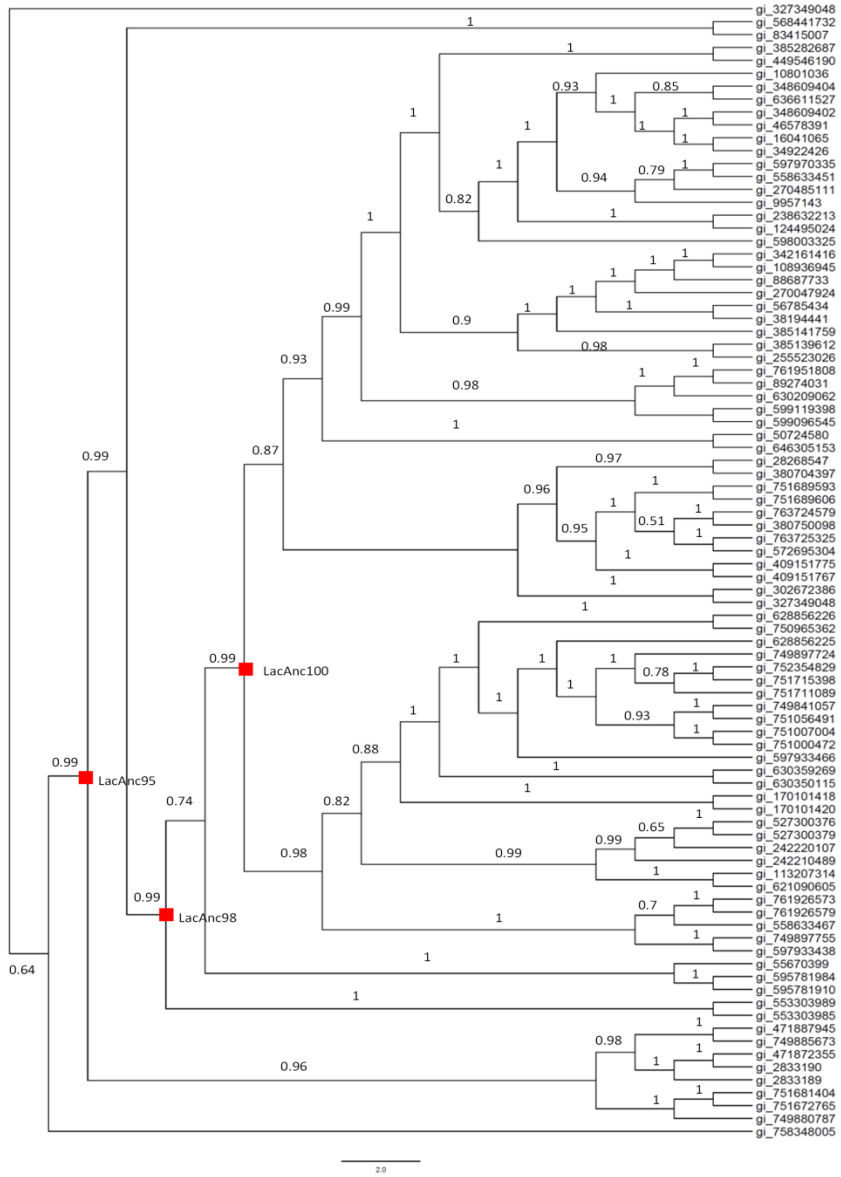


Figure 3.4.1 Inferenced evolutionary trees with the posterior probability at each node

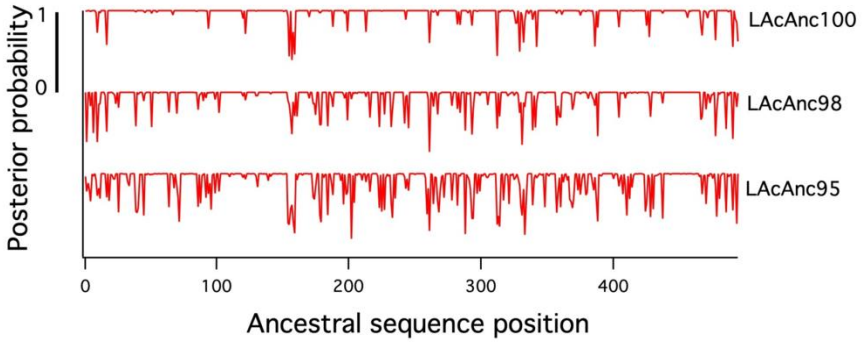


Figure 3.4.2 Robustness of the reconstructed ancestors as measure of the distribution of posterior probability

Synthesis of genes

Laccase ancestor node 95 (LacAnc95), Laccase ancestor node 98 (LacAnc98) and Laccase ancestor node 100 (LacAnc100) were synthesized by ATG biosynthetics (Merzhausen, Germany).

Protein modelling

The structural model of wild-type PM1 laccase at a resolution of 2.49 Å (1 Å=0.1 nm) (PDB code 5ANH) was used to map the mutations and to calculate distances. LacAnc100 was modelled by the Phyre2 server (protein Homoloy/analogY Recognition Engine V 2.0) (Kelley and Sternberg 2009) available at www.sbg.bio.ic.ac.uk/phyre2.

3.4.2 Directed evolution

Library creation methods

General aspects. All PCR were carried out using a gradient thermocycler (Mycycler, Biorad, USA). PCR fragments were cleaned, concentrated and loaded onto a low melting point preparative agarose gel (Bio-Rad, Hercules, CA), and then purified using the Zymoclean gel DNA recovery kit (Zymo Research, Orange, CA). PCR products were then mixed with the previously linearized vector pJRoC30 with XhoI and BamHI (at a PCR product/linearized plasmid ratio of 4:1) and transformed into competent *S. cerevisiae* cells. The whole gene was reassembled *in vivo* by transformation into *S. cerevisiae* via the design of ~40-base pairs (bp) overhangs flanking each recombination area.

Fusion of ancestral laccase with different signal: LacAnc95, LacAnc98 and LacAnc100 mature proteins were fused to the signal peptides α factor prepro-leader from PM1L evolution, the α factor prepro-leader from *Pycnoporus cinnabarinus* laccase (PcL) evolution and the chimeric preαproK from Aryl alcohol oxidase (AAO) evolution (Mate et al. 2010, Camarero et al. 2012, Viña-Gonzalez et al. 2015). To fuse LacAnc95, LacAnc98 and LacAnc100 to $\alpha^{PM1}$and α^{PcL}, the laccase gen was amplified from pUC derivative using oligonucleotide sense *OB1_95F*, *OB1_98F* and *OB1_100F* (the overhang for pJRoC30 is underlined) for α^{PM1} and *3PO_95F*, *3PO_98F* and *3PO_100F* (the overhang for pJRoC30 is underlined) for α^{PcL} and oligonucleotide antisense *RMLC* in all the cases. The α-factor-OB1 (α^{PM1}) (89 residues, including the STE13 cleavage site EAEA) was obtained from pJRoC30-OB-1 using oligonucleotide sense *RMLN* primer and oligonucleotide antisense *OB1_95R, OB1_98R and OB1_100R* (the overhang for laccases genes are underlined). The α-factor- 3PO (α^{PcL}) (89 residues, including the STE13 cleavage site EAEA) was obtained

from pJRoC30-3PO using oligonucleotide sense *RMLN* and oligonucleotide antisense *3PO_95R, 3PO_98R and 3PO_100R* (the overhang for laccases genes are underlined). The three nodes sequences were synthetized in pUC derivative with the signal peptide preαproK-AAO. In order to move these constructions to the pJRoC30, oligonucleotide sense *RMLN* and oligonucleotide antisense *RMLC* were used, **Tables 3.4.2 and 3.4.3**. The amplification reactions were carried out in a thermal cycler Mycycler (Bio-Rad). PCRs were performed in a final volume of 50 µL containing 250 nM of each primer, 10 ng of template, dNTPs at 200 µM each, 3% (vol/vol) dimethyl sulfoxide (DMSO), and 0.02 U/µL of iProof high-fidelity polymerase. The PCR cycles were as follows for the mature gene amplifications: 98°C for 30 s (1 cycle); 98°C for 10 s, 50°C for 25 s, and 72°C for 60 s (28 cycles); and 72°C for 10 min (1 cycle). The PCR cycles were as follows for the signal peptides amplifications: 98°C for 30 s (1 cycle); 98°C for 10 s, 45°C for 30 s, and 72°C for 15 s (28 cycles); and 72°C for 10 min (1 cycle).

Table 3.4.2 List of primers for fusion with α^{PM1} signal peptide

α^{PM1} LacAnc95	
OB1_95F	5'- GAGAGACTGAAGCTGAATTCGCAGCGACTGTCGTTTATGA-3'
RMLC	5'- GGGAGGGCGTGAATGTAAGC-3'
OB1_95R	5'- TCATAAACGACAGTCGCTGCGAATTCAGCTTCAGTCTCTC-3'
RMLN	5'- CCTCTATACTTTAACGTCAAGG -3'
α^{PM1} LacAnc98	
OB1_98F	5'- GAGAGACTGAAGCTGAATTCGCAGCAACCGTTGTGTATGA-3'
RMLC	5'- GGGAGGGCGTGAATGTAAGC-3'
OB1_98R	5'- TCATACACAACGGTTGCTGCGAATTCAGCTTCAGTCTCTC-3'
RMLN	5'- CCTCTATACTTTAACGTCAAGG -3'
α^{PM1}LacAnc100	
OB1_100F	5'-GAGAGACTGAAGCTGAATTCGCGATAGGACCTGTGGCAGA-3'
RMLC	5'- GGGAGGGCGTGAATGTAAGC-3'
OB1_100R	5'- TCTGCCACAGGTCCTATCGCGAATTCAGCTTCAGTCTCTC-3'
RMLN	5'- CCTCTATACTTTAACGTCAAGG -3'

Table 3.4.3 List of primers for fusion with α^{PcL} signal peptide

α^{PcL}LacAnc95	
3PO_95F	5'- GAGGGGCTGAAGCTGAATTCGCAGCGACTGTCGTTTATGA-3'
RMLC	5'- GGGAGGGCGTGAATGTAAGC-3'
3PO_95R	5'- TCATAAACGACAGTCGCTGCGAATTCAGCTTCAGCCCCTC-3'
RMLN	5'- CCTCTATACTTTAACGTCAAGG -3'
α^{PcL}LacAnc98	
3PO_98F	5'- GAGGGGCTGAAGCTGAATTCGCAGCAACCGTTGTGTATGA-3'
RMLC	5'- GGGAGGGCGTGAATGTAAGC-3'
3PO_98R	5'- TCATACACAACGGTTGCTGCGAATTCAGCTTCAGCCCCTC-3'
RMLN	5'- CCTCTATACTTTAACGTCAAGG -3'
α^{PcL}LacAnc100	
3PO_100F	5'- GAGGGGCTGAAGCTGAATTCGCGATAGGACCTGTGGCAGA-3'
RMLC	5'- GGGAGGGCGTGAATGTAAGC-3'
3PO_100R	5'- TCTGCCACAGGTCCTATCGCGAATTCAGCTTCAGCCCCTC-3'
RMLN	5'- CCTCTATACTTTAACGTCAAGG -3'

Combinatorial saturation mutagenesis CSM (162Ala, 165Val and 265Ile): Two libraries were designed in this experiment. The first saturated positions 162 and 265 and the second saturated positions 165 and 265. For accomplish these libraries 22c-trick methodology was employed (Kille et al. 2012). Therefore for the first library to saturate position 162, the oligonucleotides sense *162A-F*, *162B-F 162C-F* and the oligonucleotides antisense *162A-R, 162B-R 162C-R* were used; to saturate position 265 the oligonucleotides sense *265A-F 265B-F 265C-F* and the oligonucleotides antisense *265A-R, 265B-R 265C-R* were used; the alternative codons are underlined. For the second library to saturate position 165 the oligonucleotides sense *165A-F, 165B-F 165C-F* and the oligonucleotides antisense *165A-R*, *165B-R 165C-R* were used; to saturate position 265 the oligonucleotides sense *265A-F*, *265B-F*, *265C-F* and the oligonucleotides antisense *265A-R*, *265B-R* and *265C-R* were used; the alternative codons are underlined, **Table 3.4.4**. PCRs were performed in a final volume of 50 µL containing 250 nM of each primer, 10 ng of template, dNTPs at 200 µM each, 3% (vol/vol) dimethyl sulfoxide (DMSO), and 0.02 U/µL of iProof high-fidelity polymerase. The PCR cycles were as follows: 98°C for 30 s (1 cycle); 98°C for 10 s, 50°C for 25 s, and 72°C for 60 s (28 cycles); and 72°C for 10 min (1 cycle). 200 ng of PCR product were transformed with linearized plasmid (100 ng) into chemically competent cells. PCR products from each library were transformed independently. Transformed cells were plated

on SC drop-out plates and incubated for 3 days at 30°C. Colonies containing the whole autonomously replicating vector were exposed to high-throughput screening.

Table 3.4.4 List of primers used in CSM

Position 163	
162A-F	5'- GCTCCCCAGTTGGCCGCTNDTCCGCCCGTCCCAGATAGCAC -3'
162B-F	5'- GCTCCCCAGTTGGCCGCTVGHCCGCCCGTCCCAGATAGCAC -3'
162C-F	5'- GCTCCCCAGTTGGCCGCTTGGCCGCCCGTCCCAGATAGCAC -3'
162A-R	5'- GTGCTATCTGGGACGGGCGGAHNAGCGGCCAACTGGGGAGC -3'
162B-R	5'- GTGCTATCTGGGACGGGCGGDCBAGCGGCCAACTGGGGAGC -3'
162C-R	5'- GTGCTATCTGGGACGGGCGGCCAAGCGGCCAACTGGGGAGC -3'
Position 165	
165A-F	5'- CCAGTTGGCCGCTGCGCCGCCCNDTCCAGATAGCACCCTA -3'
165B-F	5'- CCAGTTGGCCGCTGCGCCGCCCVGHCCAGATAGCACCCTA -3'
165C-F	5'- CCAGTTGGCCGCTGCGCCGCCCTGGCCAGATAGCACCCTA -3'
165A-R	5'- TAGGGTGCTATCTGGAHNGGGCGGCGCAGCGGCCAACTGG -3'
165B-R	5'- TAGGGTGCTATCTGGDCBGGGCGGCGCAGCGGCCAACTGG -3'
165C-R	5'- TAGGGTGCTATCTGGCCAGGGCGGCGCAGCGGCCAACTGG -3'
Position 265	
265A-F	5'- GATAAGAGCGAATCCAAATNDTGGAACGACAGGATTCGCA -3'
265B-F	5'- GATAAGAGCGAATCCAAATVGHGGAACGACAGGATTCGCA -3'
265C-F	5'- GATAAGAGCGAATCCAAATTGGGGAACGACAGGATTCGCA -3'
265A-R	5'- TGCGAATCCTGTCGTTCCAHNATTTGGATTCGCTCTTATC -3'
265B-R	5'- TGCGAATCCTGTCGTTCCDCBATTTGGATTCGCTCTTATC -3'
265C-R	5'- TGCGAATCCTGTCGTTCCCCAATTTGGATTCGCTCTTATC -3'

Site Directed Recombination (SDR) library: The method applied in (Viña-Gonzalez et al. 2019) was adapted with several modifications. LacAnc100 was used as DNA template and three positions were selected as targets (A162R, V165R and I265F) for the *in vivo* site directed recombination of the wild type versus the mutated amino acid founded in the saturation mutagenesis libraries. PCR1 used oligo sense RMLN and a mix v/v of oligos antisense *162165wtRev* and *162165mutRev* (containing positions 162 and 165 underlined). PCR2 was performed with a mix v/v of oligos sense *162165wtFor* and *162165mutFor* (containing positions 162 and 165 underlined) and oligo antisense *265Rev* (containing position 265 underlined). PCR3 used oligo sense *265For* (containing position 265 underlined) and oligo antisense RMLC. **Figure 3.4.3; Table 3.4.5**. PCRs were performed in a final volume of 50 µL containing 250 nM of each primer, 10 ng of template, dNTPs at 200 µM each, 3% (vol/vol) dimethyl sulfoxide (DMSO), and 0.02 U/µL of iProof high-

fidelity polymerase. The PCR cycles were as follows: 98°C for 30 s (1 cycle); 98°C for 10 s, 50°C for 25 s, and 72°C for 60 s (28 cycles); and 72°C for 10 min (1 cycle). PCR products were mixed in equimolar amounts, 200 ng from each of the three PCRs products and transformed with linearized plasmid (150 ng) into chemically competent cells. All PCR products were transformed combined to allow *in vivo* DNA shuffling. Transformed cells were plated on SC drop-out plates and incubated for 3 days at 30°C. Colonies containing the whole autonomously replicating vector were subjected to high-throughput screening.

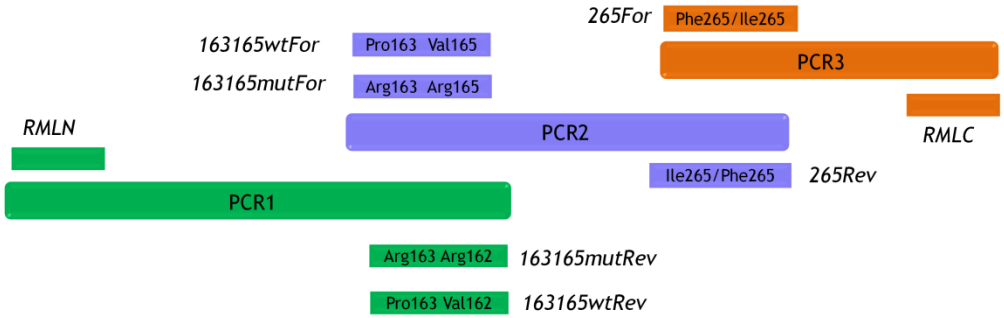


Figure 3.4.3 Primers designed for the PCR amplifications of the selected mutated positions (in black) in the site-directed recombination experiments to shuffle the best clones from saturation mutagenesis library.

Table 3.4.5 List of primers used for SDR

PCR1	
RMLN	5' CCTCTATACTTTAACGTCAAGG 3'
163165wtRev	5' GTGCTATCTGGGACGGGCGGCGCAGCGGCCAACTGGGGAGC 3'
163165mutRev	5' GTGCTATCTGGACGGGGCGGACGAGCGGCCAACTGGGGAGC 3'
PCR2	
163165wtFor	5' GCTCCCCAGTTGGCCGCTGCGCCGCCCGTCCCAGATAGCAC 3'
163165mutFor	5' GCTCCCCAGTTGGCCGCTCGTCCGCCCCGTCCAGATAGCAC 3'
265Rev	5' TGCGAATCCTGTCGTTCCRAWATTTGGATTCGCTCTTATC 3'
PCR2	
265For	5' GATAAGAGCGAATCCAAATWTYGGAACGACAGGATTCGCA 3'
RMLC	5' GGGAGGGCGTGAATGTAAGC 3'

High-throughput dual screening

Individual clones were picked and inoculated in sterile 96-well plates (Greiner Bio-One, GmbH, Germany), referred to as master plates, containing 200 µL of SEM per well. In addition to the regular SEM, the SEM without ethanol supply was also used to test the expression of the different fusion genes and signal peptides. In each plate, column number 6 was inoculated with the parent type, and one well (H1-control) was inoculated with *S. cerevisiae* transformed with pJRoC30-AAO plasmid (aryl-alcohol oxidase without activity vs

2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) or 1,3-cyclopentanedione). Plates were sealed with parafilm to prevent evaporation and incubated at 30°C, 220 rpm and 80% relative humidity in a humidity shaker (Minitron, Infors, Switzerland) for three days. Alternatively, 20 and 25°C were also tested to evaluate the expression levels of the different fusion genes and signal peptides The master plates were centrifuged (Eppendorf 5810R centrifuge, Germany) for 10 min at 2,500 *g* and 4°C. Aliquots of the supernatants were transferred from the master plates to two replica plates with the aid of a liquid handler robotic station Freedom EVO (Tecan, Switzerland), 80 µL of mixture for β-diketone screening and 20 µL to the ABTS screening. To estimate the β-diketone activity 120 µL of 100 mM citrate phosphate buffer pH 5.0 containing 50 mM 1,3-cyclopentanedione were added to the replica plates containing 80 µL of supernatant. Plates were stirred briefly and the absorption at 450 nm was measured in end point mode in the plate reader (SPECTRAMax Plus 384, Molecular Devices, Sunnyvale, CA). Afterwards the plates were sealed with parafilm and incubated in shaker at 35°C. Incubation took place for 40 hours and then plates were measured again. The activities were calculated from the difference between the absorption after incubation and that of the initial measurement normalized against the parental type in the corresponding plate. To estimate the ABTS activity 180 µL of 100 mM citrate phosphate buffer pH 4.0 containing 1 mM ABTS were added to the replica plates containing 20 µL of supernatant. Plates were stirred briefly and the absorption at 418 nm (εABTS $^{\bullet+}$ = 36,000 M^{-1} cm^{-1}) was followed in kinetic mode in the plate reader. To rule out false positives, two consecutive rescreenings were carried out according to the protocol previously reported with some modifications (Mate et al. 2010).

DNA sequencing

Plasmids containing variants HRPLs were sequenced by GATC-Biotech. The samples were prepared with 5 µL of 100 ng/µL plasmid and 5 µL of 5 µM of each primer, and were as follows: *RMLN*, *FS*, *RS* and *RMLC* for extant laccase and RMLN, LacAnc100Forward, LacAnc100Foward2 and RMLC for ancestral laccase and mutants, **Table 3.4.6**.

Table 3.4.6 List of primers used for sequencing

OB-1	
RMLN	5' CCTCTATACTTTAACGTCAAGG 3'
FS	5' ACGACTTCCAGGTCCCTGACCAAGC 3'
RS	5' TCAATGTCCGCGTTCGCAGGGA 3'
RMLC	5' GGGAGGGCGTGAATGTAAGC 3'
LacAnc100	
RMLN	5' CCTCTATACTTTAACGTCAAGG 3'
Lac100Forward	5' CGTTAGCCGTGATAAACGTAGAGCAGGGAAA 3'
Lac100Foward2	5' GATTCTTTCAACACGGGACAAATTGGGCTGACG 3'
RMLC	5' GGGAGGGCGTGAATGTAAGC 3'

3.4.3 Biochemical characterization

Production and purification

A single colony from the *S. cerevisiae* clone containing the parent or mutant gene was picked from a SC drop-out plate, inoculated in minimal medium (10 mL) and incubated for 48 h at 30°C and 230 rpm. An aliquot of cells was removed and used to inoculate minimal medium (100 mL) in a 500 mL flask (at OD600 of 0.25). The cells completed two growth phases (6–8 h) and then expression medium (900 mL) was inoculated with the pre-culture (100 mL) (OD600 of 0.1) in a 1 L flask. After incubating for 72 h at 30°C and 230 rpm (maximal laccase activity; OD600 = 25-30), the cells were recovered by centrifugation at 4,500 rpm (4°C) and the supernatant was double-filtered (using both glass membrane and a nitrocellulose membrane of 0.45 µm pore size). The purification protocol was adapted from (Mate, Gonzalez-Perez, Kittl, et al. 2013) with several modifications. Recombinant laccase purification was achieved by anion exchange chromatography (ÄKTA purifier, GE Healthcare, WI, US). The crude extract was concentrated and dialyzed in bis-tris 20 mM buffer at pH 6.0 (buffer A) by tangential ultrafiltration (Pellicon; Millipore, Temecula, CA, US) through a 10-kDa-pore-size membrane (Millipore) by means of a peristaltic pump (Masterflex Easy Load; Cole-Parmer, Vernon Hills, IL). The sample was filtered and loaded onto a strong anion-exchange column (HiTraP QFF, Amersham Bioscience) pre-equilibrated with buffer A. The proteins were eluted with a linear gradient from 0 to 1 M of NaCl in two phases at a flow rate of 1 mL/min: from 0 to 50% over 40 min and from 50 to 100% over 10 min. Fractions with laccase activity were pooled, concentrated, dialyzed against buffer A and further purified by HPLC- Photodiode array (PDA) coupled with a 10 µm high resolution anion exchange

Biosuite Q (Waters, MA, USA) pre-equilibrated with buffer A. The proteins were eluted with a linear gradient from 0 to 1 M of NaCl at a flow rate of 1 mL/min in two phases: from 0 to 40% in 50 min and from 40 to 100% in 10 min. The fractions with laccase activity were pooled, dialyzed against 20 mM bis-tris buffer pH 6.0, concentrated and stored at 4°C.

pH activity profile

Appropriate dilutions of supernatants were prepared with the help of the robot in such a way that aliquots of 20 µL give rise to a linear response in kinetic mode. Vessels containing 100 mM Britton and Robinson buffer with 1 mM of ABTS, 5 mM of DMP, 5 mM of $K_4[Mo(CN)_8]$ and 50 mM of Guaiacol were prepared at pH values of 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, and 9.0. The assay started when the different reactions mixtures of ABTS, DMP, $K_4[Mo(CN)_8]$ and Guaiacol was added to each well. The activities were measured in triplicate in kinetic mode, and the relative activity (in percent) is based on the maximum activity for each variant in the assay.

pH stability

To measure pH stability, pure enzyme samples were incubated at different times over a range of pH values in 100 mM Britton and Robinson buffer (2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, and 9.0). Samples were removed at different times (0, 4, 24, 48, 72 and 144 hours screening. The activities were measured in triplicate in kinetic mode, and the residual activity was deduced from the activity obtained at each time.

Half life *($t_{1/2}$)*

Appropriate dilutions of supernatants were prepared (in pH 3.0 and 6.0) with the help of the robot in such a way that aliquots of 20 µL give rise to a linear response in kinetic mode. 50 µL (from both selected mutants and parent types) were used for each point in the incubation. A fix temperature of 70°C and 50°C was established, for pH 6.0 and 3.0, respectively. Samples were removed at different times (in min): 0, 5, 15, 25, 35, 55, 75, 95 and 120 minutes) from the thermocycler (Mycycler, Biorad, US) and chilled out on ice for 10 min. After that, samples of 20 µL were removed and incubated at room temperature for 5 min. Finally, samples were subjected to the same ABTS-based colorimetric assay described above for the screening. Thermostabilities values were deduced from the ratio between the residual activities incubated at different temperature points and the initial activity at room temperature.

Steady-state kinetic constants

ABTS, DMP, sinapic acid and guaiacol kinetics constants for pure laccases were estimated in sodium phosphate/citrate buffer (pH 4.0, 100 mM for ABTS and pH 5.0, 100 mM for DMP, sinapic acid and guaiacol). Reactions were performed in triplicate, and substrate oxidations were followed through spectrophotometric changes (ε_{418}ABTS=36,000 M^{-1} cm^{-1}; ε_{469} DMP=27,500 M^{-1} cm^{-1}; ε_{512} Sinapic acid=14,066 M^{-1} cm^{-1} and ε_{470} Guaiacol=12,100 M^{-1} cm^{-1}). To calculate the K_m and k_{cat} values, the average V_{max} was represented against substrate concentration and fitted to a single rectangular hyperbola function with SigmaPlot 10.0, where parameter *a* was equal to k_{cat} and parameter *b* was equal to K_m.

3.5 Methodology employed in Chapter III: Consensus design of an evolved high-redox potential laccase

3.5.1 Consensus sequence

Multiple Sequence Alignments (MSAs)

Fungal laccases homologues were retrieved from the April 2015 release of the complete genomes database available at National Center for Biotechnology Information (NCBI) (Coordinators 2018). Three different MSA were created; MSA1 sequences were obtained from NCBI database upon Basic Local Alignment Search Tool for proteins (BLASTp) search (Altschul et al. 1997), using the following parameters: PM1L as query sequence, non-redundant protein sequences, maximum of 250 sequences to display and a threshold of percentage identity between 35% and 98%; MSA2 sequences were retrieved following the same steps as for MSA1 sequences but establishing a maximum of 1000 sequences to display; For MSA3 sequences, all the sequences of basidiomycete laccases were recovered from the Laccase and Multicopper Oxidase Engineering Database (LccED) (Sirim et al. 2011). All sequences including terms such as "hypothetical", "predicted" or "putative" were discarded, as well of those sequences coming from insects or ascomycetes, resulting in a set of 235 for MSA1, 819 for MSA2 and 477 for MSA3. PM1 sequence was added to the 3 groups of sequences and then they were aligned independently using MUSCLE (Edgar 2004) generating the 3 MSAs as a result.

Data Mining from the MSAs

The 3 MSAs were imported into Microsoft Excel to refine them as described elsewhere (Durani and Magliery 2013) with minor modifications. Relative Entropy (RE, **Figure 3.5.1**) metric was calculated for each position, using the yeast proteome codon usage as reference distribution. Mutual information (MI, **Figure 3.5.2**) metric was calculated as well as the MI noise from a randomized version of the alignment to avoid MI values with no statistical significance. Those Mi values below the MI noise were discarded. The outcoming consensus sequence was compared with OB-1 in order to localize the possible consensus mutations. The thresholds employed in RE and MI for sieving consensus positions are indicated in the results and discussion from Chapter III.

$$RE = \sum_i p\ (xi)\ \log\frac{p(x_i)}{q(x_i)}$$

Figure 3.5.1 Relative Entropy formula. Where $p(x_i)$ is the probability of observing residue i at position x in the MSA, and $q(x_i)$ is the probability of observing residue i in the yeast proteome. *S.cerevisiae* amino acid codon usage is as follows: A-5.6%, C-1.3%, D-5.8%, E-6.5%, F-4.4%, G-5.1%, H-2.1%, I-6.5%, K-7.3%, L-9.5%, M-2.1%, N-6.1%, P-4.3%, Q-3.9%, R-4.4%, S-8.9%, T-5.9%, V-5.7%, W-1.0%, Y-3.4%.

$$MI_{xy} = \sum_i \sum_j p(x_{i,}\, y_i)\log\frac{p(x_i, y_i)}{p(x_i)p(y_i)}$$

Figure 3.5.2 Mutual Information formula. Where $p(x_i)$ is the probability of observing residue i at position x in the MSA, $p(y_i)$ is the probability of observing residue j at position y, and $p(x_i)p(y_i)$ in the joint probability of observing residue i at position x and residue j at position y. This value is summed over all the 400 possible pairs of amino acid residues

Protein modelling

The structural model of wild-type PM1L at a resolution of 2.49 Å (1 Å=0.1 nm) (PDB code 5ANH) was used to map the consensus mutations using PyMOL Molecular Graphics System (Version 1.6 Schrödinger, LLC).

3.5.2 Consensus design

Library creation methods

General aspects. PCR fragments were cleaned, concentrated and loaded onto a low melting point preparative agarose gel (Bio-Rad, Hercules, CA), and then purified using the Zymoclean gel DNA recovery kit (Zymo Research, Orange, CA). PCR products were then mixed with the previously linearized vector pJRoC30 with XhoI and BamHI (at a PCR product/linearized plasmid ratio of 4:1) and transformed into competent *S. cerevisiae* cells. The whole gene was reassembled *in vivo* by transformation into *S. cerevisiae* via the design of ~40-bp overhangs flanking each recombination area.

Site Directed Mutagenesis: The 20 single mutants from the consensus analysis were designed by *In vivo* Overlap Extension (IVOE) (Alcalde 2010). OB-1 mutant was used as template and two independent PCRs (PCR1 and PCR2) were performed for each of the 20 constructions, **Table 3.5.1**. Reaction mixtures were prepared in a final volume of 50 µL containing DNA template (0.2 ng/µL), 0.25 µM of each primer, 0.3 mM dNTPs (0.075 mM each), 3% (v:v) dimethylsulfoxide (DMSO) and 0.05 U/µL Pfu DNA polymerase.

PCRs were performed in a thermocycler (Mycycler, Bio-Rad,Hercules, CA, USA) and parameters were: 94°C for 2 min (1 cycle); 94°C for 45 s, 50°C for 30 s, 72°C for 1 min (30 cycles); and 72°C for 10 min (1 cycle). PCR products were mixed in equimolar amounts, 200 ng PCR1 product and 200 ng PCR2 product and transformed with linearized plasmid (100 ng) into chemically competent cells. PCR products were transformed independently in pairs, so just one position was mutated per transformation. Transformed cells were plated on SC drop-out plates and incubated for 3 days at 30°C. Colonies containing the whole autonomously replicating vector were picked and after analysed 16 clones per mutation, the selected plasmids were extracted and sequenced to verify the site directed mutagenesis. Once the single mutants were confirmed they were submitted to the third rescreening ($t_{1/2}$ determination).

Table 3.5.1 List of primers used for site directed mutagenesis

I25V	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 25R: 5' GGGGAAGACGTCGTTGACCAGGACGGCCTGCCGAGAGAAACCATC 3'
	PCR 2	25F: 5' GATGGTTTCTCTCGGCAGGCCGTCCTGGTCAACGACGTCTTCCCC 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
V27A	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 27R : 5' GGACTGGGGAAGACGTCGTTGGCCAGGATGGCCTGCCGAGAGA 3'
	PCR 2	27F: 5' TCTCTCGGCAGGCCATCCTGGCCAACGACGTCTTCCCCAGTCC 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
S33G	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 33R: 5' TTGTTCCCCGTAATGAGGGGACCGGGGAAGACGTCGTTGACCAG 3'
	PCR 2	33F: 5' CTGGTCAACGACGTCTTCCCCGGTCCCCTCATTACGGGGAACAA 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
D50N	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 50R : 5' ATGGTGTGGTTGGTCATGTTGTTGATGACATTGAGTTGGAAACG 3'
	PCR 2	50F: 5' CGTTTCCAACTCAATGTCATCAACAACATGACCAACCACACCAT 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
M52L	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 52R: 5' TTCAACATGGTGTGGTTGGTCAGGTTGTCGATGACATTGAGTTG 3'
	PCR 2	52F: 5' CAACTCAATGTCATCGACAACCTGACCAACCACACCATGTTGAA 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
A92S	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 92R: 5' ACCTGGAAGTCGTAAAGGAACGAATGCCCGGTAGAAATCGGGCA 3'
	PCR 2	92F: 5' TGCCCGATTTCTACCGGGCATTCGTTCCTTTACGACTTCCAGGT 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
D142E	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 142R: 5' CGCGAGAGTGATTACAGTGGACTCGTCATCAACATCGTAAAGGC 3'
	PCR 2	142F: 5' GCCTTTACGATGTTGATGACGAGTCCACTGTAATCACTCTCGCG 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
T179G	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 179R: 5' TGACAGCCAAATCGGCGTTGAGTCCGTTGATGCTGCGACCGAGGCCG 3'
	PCR 2	179F: 5' CGGCCTCGGTCGCAGCATCAACGGACTCAACGCCGATTTGGCTGTCA 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
Y208F	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 208R: 5' GAGTGACCATCAATGCTGAACGAATAATTCGGGTCGCATGACAGCG 3'
	PCR 2	208F: 5' CGCTGTCATGCGACCCGAATTTCACGTTCAGCATTGATGGTCACTC 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
L217M	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 217R: 5' GCCGTCCGCCTCGATGACGGTCATAGAGTGACCATCAATGCTGAA 3'
	PCR 2	217F: 5' TTCAGCATTGATGGTCACTCTATGACCGTCATCGAGGCGGACGGC 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
V219I	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 219R: 5' TTCACGCCGTCCGCCTCGATGATGGTCAGAGAGTGACCATCAAT 3'
	PCR 2	219F: 5' ATTGATGGTCACTCTCTGACCATCATCGAGGCGGACGGCGTGAA 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'

Q230L	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 230R: 5' GATCTGGATGGAGTCGACAGTCAGGGGCTTGAGATTCACGCCGTC 3'
	PCR 2	230F: 5' GACGGCGTGAATCTCAAGCCCCTGACTGTCGACTCCATCCAGATC 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
A240G	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 240R: 5' GAGCACAAACGAGTACCGCTGGCCAGGGAAGATCTGGATGGAGT 3'
	PCR 2	240F: 5' ACTCCATCCAGATCTTCCCTGGCCAGCGGTACTCGTTTGTGCTC 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
D250N	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 250R: 5' CAGTAGTTGTCCACATCCTGATTTGCGTTGAGCACAAACGAGTA 3'
	PCR 2	250F: 5' TACTCGTTTGTGCTCAACGCAAATCAGGATGTGGACAACTACTG 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
N268G	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 268R: 5' GGAGTTAACGCCGCCGTCGAAGCCCCTGGTCCCGGAGTTGGGAAG 3'
	PCR 2	268F: 5' CTTCCCAACTCCGGGACCAGGGGCTTCGACGGCGGCGTTAACTCC 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
T306P	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 306R: 5' GGCGCAGCGGTGCCTTCGAGAGGGGTCAGGGCGGACTCCACCAA 3'
	PCR 2	306F: 5' TTGGTGGAGTCCGCCCTGACCCCTCTCGAAGGCACCGCTGCGCC 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
M327L	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 327R: 5' CCGCCGGCAAAGCCGAAAGCCAAGTTGAGAGCCAGGTCGACACC 3'
	PCR 2	327F: 5' GGTGTCGACCTGGCTCTCAACTTGGCTTTCGGCTTTGCCGGCGG 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
Y421R	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 421R: 5' AGCCCGTGTTGACGACGTCGCCGTAGACCGGGTTCGCGTAGTT 3'
	PCR 2	421F: 5' TACAACTACGCGAACCCGGTCCGCCGCGACGTCGTCAACACGGG 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
F454W	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 454R: 5' CGTGAACCCAGCCTCAAGGTGCCAGTCGATGTGGCAGTGGAGGA 3'
	PCR 2	454F: 5' TCCTCCACTGCCACATCGACTGGCACCTTGAGGCTGGGTTCACG 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'
M464F	***PCR 1***	RMLN: 5' CCTCTATACTTTAACGTCAAGG 3' 464R: 5' GACGTCGGGAATGTCCTCGGCGAAGACGACCGTGAACCCAGCCTC 3'
	PCR 2	464F: 5' GAGGCTGGGTTCACGGTCGTCTTCGCCGAGGACATTCCCGACGTC 3' RMLC: 5' GGGAGGGCGTGAATGTAAGC 3'

Site Directed Recombination (SDR) library: The method applied in (Viña-Gonzalez et al. 2019) was adapted with several modifications. DooKu variant (A240G) was used as template and mutations M52L, T179G, L217M, V219I, Q230L and D250N were targeted for SDR *in vivo*. PCR1 used oligo sense RMLN and oligo antisense *52Rev* (containing position 52). PCR2 was performed with oligo sense *52For* (containing position 52) and a mix v/v of oligos antisense *179wRev* and *179mutRev* (containing position 179). PCR3 used v/v of oligos sense *179wFor* and *179mutFor* (containing position 179) and oligo antisense *219Rev* (containing positions 217, 219 and 230). PCR4 used oligo sense *219For* (containing positions 217, 219 and 230) and oligo antisense *250Rev* (containing position 250). PCR5 was performed with oligo sense *250For* (containing position 250) and oligo antisense RMLC, **Figure 3.5.3; Table 3.5.2**. For the *in vivo* assembly of the whole gene, the fragments were designed with overhangs of around 40 bp flanking them. Reaction mixtures were prepared in a final volume of 50 µL containing DNA template (0.2 ng/µL), 0.25 µM of each primer, 0.3 mM dNTPs (0.075 mM each), 3% (v:v) dimethylsulfoxide (DMSO) and 0.05 U/µL Pfu DNA polymerase. The amplification parameters were 94°C for 2 min (1

cycle); 94°C for 45 s, 50°C for 30 s, 72°C for 2 min (30 cycles); and 72°C for 10 min (1 cycle). PCR products were mixed in equimolar amounts, 240 ng from each of the five PCRs products and transformed with linearized plasmid (300 ng) into chemically competent cells. All PCR products were transformed combined to allow *in vivo* shuffling. Transformed cells were plated on SC drop-out plates and incubated for 3 days at 30°C. Colonies containing the whole autonomously replicating vector were submitted to high-throughput screening for thermostability.

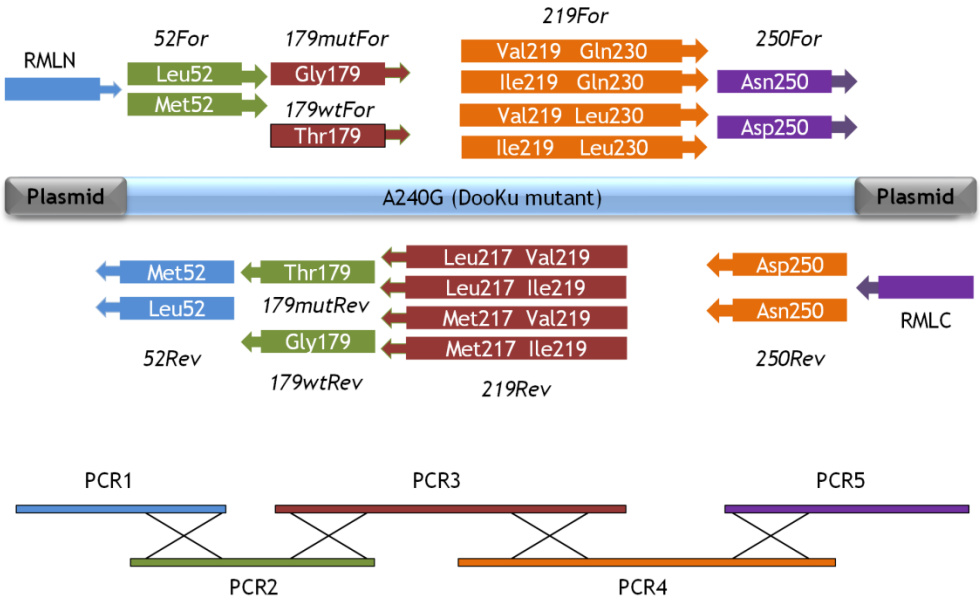


Figure 3.5.3 Site directed recombination *in vivo*. Primers designed for the PCR amplifications of the selected mutated positions (in black) in the site-directed recombination experiments to shuffle the best consensus mutations.

Table 3.5.2 List of primers used for SDR.

PCR1	
RMLN	5' CCTCTATACTTTAACGTCAAGG 3'
52Rev	5' TTCAACATGGTGTGGTTGGTCAWGTTGTCGATGACATTGAGTTG 3'
PCR2	
52For	5' CAACTCAATGTCATCGACAACWTGACCAACCACACCATGTTGAA 3'
179wtRev	5' TGACAGCCAAATCGGCGTTGAGCGTGTTGATGCTGCGACCGAGGCCG 3'
179mutRev	5' TGACAGCCAAATCGGCGTTGAGTCCGTTGATGCTGCGACCGAGGCCG 3'
PCR3	
179wtFor	5' CGGCCTCGGTCGCAGCATCAACACGCTCAACGCCGATTTGGCTGTCA 3´
179mutFor	5' CGGCCTCGGTCGCAGCATCAACGGACTCAACGCCGATTTGGCTGTCA 3'
217Rev	5' CGACAGTCWGGGGCTTGAGATTCACGCCGTCCGCCTCGATGAYGGTCAKAGAGTGACCA 3'
PCR4	
217For	5' TGGTCACTCTMTGACCRTCATCGAGGCGGACGGCGTGAATCTCAAGCCCCWGACTGTCG 3´
250Rev	5' CAGTAGTTGTCCACATCCTGATYTGCGTTGAGCACAAACGAGTA 3'
PCR5	
250For	5' TACTCGTTTGTGCTCAACGCARATCAGGATGTGGACAACTACTG 3´
RMLC	5' GGGAGGGCGTGAATGTAAGC 3'

High-throughput dual screening

Individual clones were picked and inoculated in sterile 96-well plates (Greiner Bio-One, GmbH, Germany), referred to as master plates, containing 200 µL of SEM per well. In each plate, column number 6 was inoculated with the parent type, and one well (H1-control) was inoculated with *S. cerevisiae* transformed with pJRoC30-AAO plasmid (aryl-alcohol oxidase without activity *vs.* ABTS). Plates were sealed with parafilm to prevent evaporation and incubated at 30°C, 220 rpm and 80% relative humidity in a humidity shaker (Minitron, Infors, Switzerland) for three days. The master plates were centrifuged (Eppendorf 5810R centrifuge, Germany) for 10 min at 2,500 *g* and 4°C. Aliquots of the supernatants were transferred from the master plates to two replica plates with the aid of a liquid handler robotic station Freedom EVO (Tecan, Switzerland), 50 µL of mixture to a thermocycler plate (Multiply PCR plate without skirt, neutral, Sarstedt, Germany) and 20 µL to the initial activity plate. Thermocycler plates were sealed with thermoresistant film (Deltalab, Spain) and incubated at the corresponding temperature using a thermocycler (MyCycler, Biorad, USA). Incubation took place for 10 min (so that the assessed activity was reduced 2/3 of the initial activity). Afterwards, thermocycler plates were placed on ice for 10 min and further incubated for 5 min at room temperature. 20 µL of supernatants were transferred from both thermocycler and initial activity plates to new plates to estimate

the initial activities and residual activities values by adding 180 µL of 100 mM citrate phosphate buffer pH 4.0 containing 1 mM ABTS. Plates were stirred briefly and the absorption at 418 nm (εABTS •+ = 36,000 M^{-1} cm^{-1}) was followed in kinetic mode in the plate reader (SPECTRAMax Plus 384, Molecular Devices, Sunnyvale, CA). The same experiment was performed for both the initial activity plate and residual activity plate. The values were normalized against the parent type of the corresponding plate and selected variants came from the ratio between residual activities and initial activities values, **Figure 3.5.4**. To rule out false positives, two consecutive rescreenings were carried out according to the protocol previously reported with some modifications (Mate et al. 2010). A third rescreening was also implemented for this work: Fresh transformants of selected mutants from SDR and parent types were cultivated (10 mL) in 100 mL flask following the procedure described elsewhere (Mate et al. 2010). Supernatants were analyzed in terms of $t_{1/2}$ and initial activity.

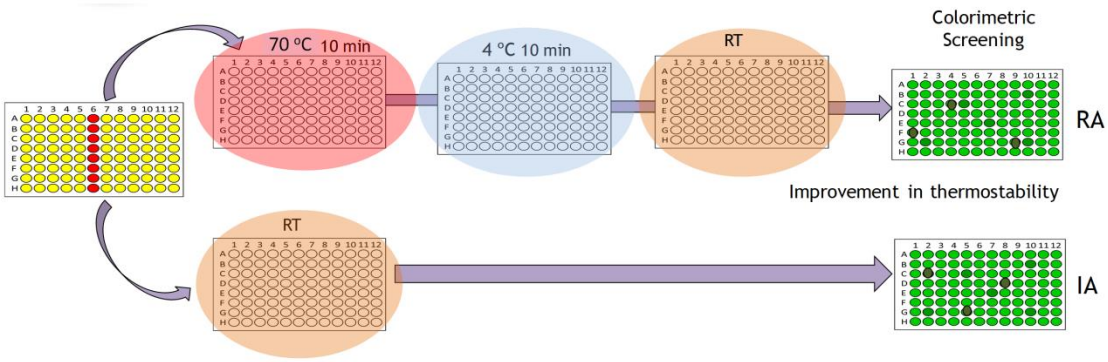


Figure 3.5.4 HTS dual screening protocol used for screening thermostability. The improvement in thermostability is calculated from the ratio between residual activities and initial activities values

DNA sequencing

Plasmids containing HRPLs variants were sequenced by GATC-Biotech. The samples were prepared with 5 µL of 100 ng/µL plasmid and 5 µL of 5 µM of each primer, and were as follows: *RMLN*, *FS*, *RS* and *RMLC*, **Table 3.5.3**.

Table 3.5.3 List of primers used for sequencing

OB-1	
RMLN	5' CCTCTATACTTTAACGTCAAGG 3'
FS	5' ACGACTTCCAGGTCCCTGACCAAGC 3'
RS	5' TCAATGTCCGCGTTCGCAGGGA 3'
RMLC	5' GGGAGGGCGTGAATGTAAGC 3'

3.5.3 Biochemical characterization

Production and purification

A single colony from the *S. cerevisiae* clone containing the parental or mutant laccase gene was picked from a SC drop-out plate, inoculated in 10 mL of minimal medium and incubated for 48 h at 30°C and 225 rpm. An aliquot of cells was removed and inoculated into a final volume of 50 mL of minimal medium in a 500 mL flask (optical density, OD_{600}=0.25). Incubation proceeded until two growth phases were completed (6 to 8 h) and then, 450 mL of expression medium was inoculated with the 50 mL preculture in a 2L flask (OD_{600}=0.1). After incubating for 72 h at 30°C and 225 rpm (maximal laccase activity; OD_{600}=28-30), the cells were separated by centrifugation for 20 min at 8000 rpm (4°C) (Avantin J-E Centrifuge, Beckman Coulter, Fullerton, CA) and the supernatant was double-filtered (through both a glass and then nitrocellulose membrane of 0.45 μm pore size). The crude extract was first submitted to a fractional precipitation with ammonium sulphate at 55 % (first cut) and the pellet was then removed before the supernatant was subjected to 75 % ammonium sulphate precipitation (second cut). The final pellet was recovered in 20 mM Bis Tris buffer pH 6.2 (buffer A) and the sample was filtered and loaded onto the fast protein liquid chromatography (FPLC) equipment (ÄKTA purifier, GE Healthcare, WI, US) coupled with a strong anionic exchange column (HiTraP QFF, Amersham Bioscience) pre-equilibrated with buffer A. The proteins were eluted with a linear gradient from 0 to 1 M of NaCl in two phases at a flow rate of 1 mL/min: from 0 to 45 % over 50 min and from 45 to 100 % over 10 min. Fractions with laccase activity were pooled, concentrated, dialyzed against buffer A and loaded onto a 10 μm high resolution anion exchange Biosuite Q (Waters, MA, USA) preequilibrated with buffer A. The proteins were eluted with a linear gradient from 0 to 1 M of NaCl at a flow rate of 1 mL/min in two phases: from 5 to 16 % in 30 min and from 16 to 100 % in 10 min. The fractions with laccase activity were pooled, dialyzed against 10 mM Bis Tris buffer pH 7.0, concentrated and stored at 4°C. The purified laccases were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). All protein concentrations were determined using the Bio-Rad protein reagent and bovine serum albumin as a standard.

pH activity profile

Appropriate dilutions of enzyme samples were prepared with the help of the liquid handler robotic station in such a way that aliquots of 20 μL give rise to a linear response in kinetic mode. Vessels containing 100 mM Britton and Robinson buffer with 1 mM of

ABTS or 3 mM of DMP, were prepared at pH values of 2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, and 9.0. The assay started when the different reactions mixtures of ABTS or DMP was added to each well. The activities were measured in triplicate in kinetic mode, and the relative activity (in percent) is based on the maximum activity for each variant in the assay.

pH stability

Enzyme samples were incubated at different times over a range of pH values in 100 mM Britton and Robinson buffer (2.0, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, and 9.0). Samples were removed at different times (0, 4, 24, 48, 72 and 144 hours) and subjected to the same ABTS-based colorimetric assay described above for the screening. The activities were measured in triplicate in kinetic mode, and the residual activity was deduced from the activity obtained at each time.

Half life *($t_{1/2}$)*

Appropriate dilutions of enzyme samples were prepared with the help of the liquid handler robotic station in such a way that aliquots of 20 µL give rise to a linear response in kinetic mode. 50 µL were used for each point in the incubation. A fix temperature of 70°C was established for consensus mutations and 75°C for the SDR variants. Samples were removed at different times (in min): 0, 5, 15, 25, 35, 55, 75, 95 and 120 minutes) from the thermocycler (Mycycler, Biorad, US) and chilled out on ice for 10 min. After that, samples of 20 µL were removed and incubated at room temperature for 5 min. Finally, samples were subjected to the same ABTS-based colorimetric assay described above for the screening. Thermostabilities values were deduced from the ratio between the residual activities incubated at different temperature points and the initial activity at room temperature.

Estimation of initial activity

Appropriate dilutions of enzyme samples were prepared with the help of the liquid handler robotic station in such a way that aliquots of 20 µL give rise to a linear response in kinetic mode. Initial activities for ABTS, DMP, sinapic acid, violuric acid and $K_4[Mo(CN)_8]$ were estimated in sodium phosphate/citrate buffer (pH 4.0, 100 mM). Reactions were performed in triplicate, and substrate oxidations were followed through spectrophotometric changes (ε_{418}ABTS=36,000 M^{-1} cm^{-1}; ε_{469} DMP=27,500 M^{-1} cm^{-1}; ε_{512} Sinapic acid=14,066 M^{-1} cm^{-1}; ε_{515} Violuric acid=98 M^{-1} cm^{-1} and ε_{388} $K_4[Mo(CN)_8]$ =1,460 M^{-1} cm^{-1}).

Steady-state kinetic constants

ABTS, DMP, sinapic acid, guaiacol and $K_4[Mo(CN)_8]$ kinetics constants were estimated in sodium phosphate/citrate buffer (pH 4.0, 100 mM). Reactions were performed in triplicate, and substrate oxidations were followed through spectrophotometric changes (ε_{418}ABTS=36,000 M^{-1} cm^{-1}; ε_{469} DMP=27,500 M^{-1} cm^{-1}; ε_{512} Sinapic acid=14,066 M^{-1} cm^{-1}; ε_{470} Guaiacol=12,100 M^{-1} cm^{-1} and ε_{388} $K_4[Mo(CN)_8]$ =1,460 M^{-1} cm^{-1}). To calculate the K_m and k_{cat} values, the average V_{max} was represented against substrate concentration and fitted to a single rectangular hyperbola function with SigmaPlot 10.0, where parameter *a* was equal to k_{cat} and parameter *b* was equal to K_m

4 RESULTS AND DISCUSSION

4.1 Chapter I: Directed *-in vitro-* evolution of Precambrian and extant Rubiscos.

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is an ancient, catalytically conserved, and low-efficient enzyme that plays a central role in the biosphere carbon cycle through the photosynthetic CO_2-fixation. The design of improved Rubiscos to increase agriculture productivity has hitherto relied on the use of biological *-in vivo-* selection systems precluding the exploration of biochemical traits that are not wired to the cell survival. We present here a directed *-in vitro-* evolution platform that put the enzyme outside of its biological context to open a new venue in Rubisco engineering. Precambrian -resurrected- and extant Rubiscos form II from *Rhodospirillum rubrum* were subjected to a broad ensemble of directed evolution strategies aimed at improving stability. The most recent common ancestor of proteobacteria -dating back to 2.4 billion years or Gigayears (Gyr) ago- showed strikingly tolerant to mutations while a panel of thermostable extant and ancestral Rubisco mutants was disclosed by adaptive evolution, focused evolution and neutral genetic drift. The conjunction of Precambrian Rubisco resurrection with cutting-edge directed *-in vitro-* evolution methods provides a reliable roadmap for designing stable and highly evolvable Rubiscos.

This chapter is based on the article: Gomez-Fernandez, B.J., Garcia-Ruiz, E., Martin-Diaz, J., Gomez de Santos, P., Santos-Moriano, P., Plou, Ballesteros, A., Garcia, M., Rodriguez, M., Risso, V.A., Sanchez-Ruiz, J.M., Whitney, S.M. and Alcalde, M. (2018) Directed *-in vitro-* evolution of Precambrian and extant Rubiscos, *Scientific Reports* 8: 5532.

4.1.1 Directed *-in vitro-* evolution of extant Rubisco

The Rubisco form II from the proteobacteria *R. rubrum* (RubRr) was chosen as template to start the directed *-in vitro-* evolution and the ancestral resurrection enterprises. This enzyme is formed by two L subunits arranged into a functional dimer -L_2-, it is heterologous expressed in *E. coli* and it does not require the helper protein rubisco activase (Portis et al. 2003). Given that Rubisco can become toxic to the *E. coli* cells, to achieved sufficient RubRr expression levels without compromising the host viability, the culture and induction conditions were optimized for the expression of RubRr in microtiter plate format (~40 Units/g of lysate from clones grown in 96 well plates). To evolve Rubisco *in vitro*, a dual HTS assay that couples both spectrophotometric and analytic detection methods was applied. The spectrophotometric method is a NADH depletion assay based

on an enzyme cascade reaction previously reported (Sulpice et al. 2007) that we modified to measure activities in RubRr mutant libraries, **Figure 4.1.1A and Figure 3.3.3 (from Materials and Methods section)**. This assay was complemented with a new HPLC-ELSD ion pair method that allowed us to measure with high accuracy the direct consumption of RuBP and the formation of 3PGA by Rubisco under CO_2 saturated conditions, **Figure 4.1.1B, C**. Both methods were validated by checking the responses and the coefficient of variation (<13%) in RubRr parental type fresh lysates from 96-well plates microcultures, **Figure 4.1.2**. Two consecutive re-screens were applied to rule out the presence of false positives.

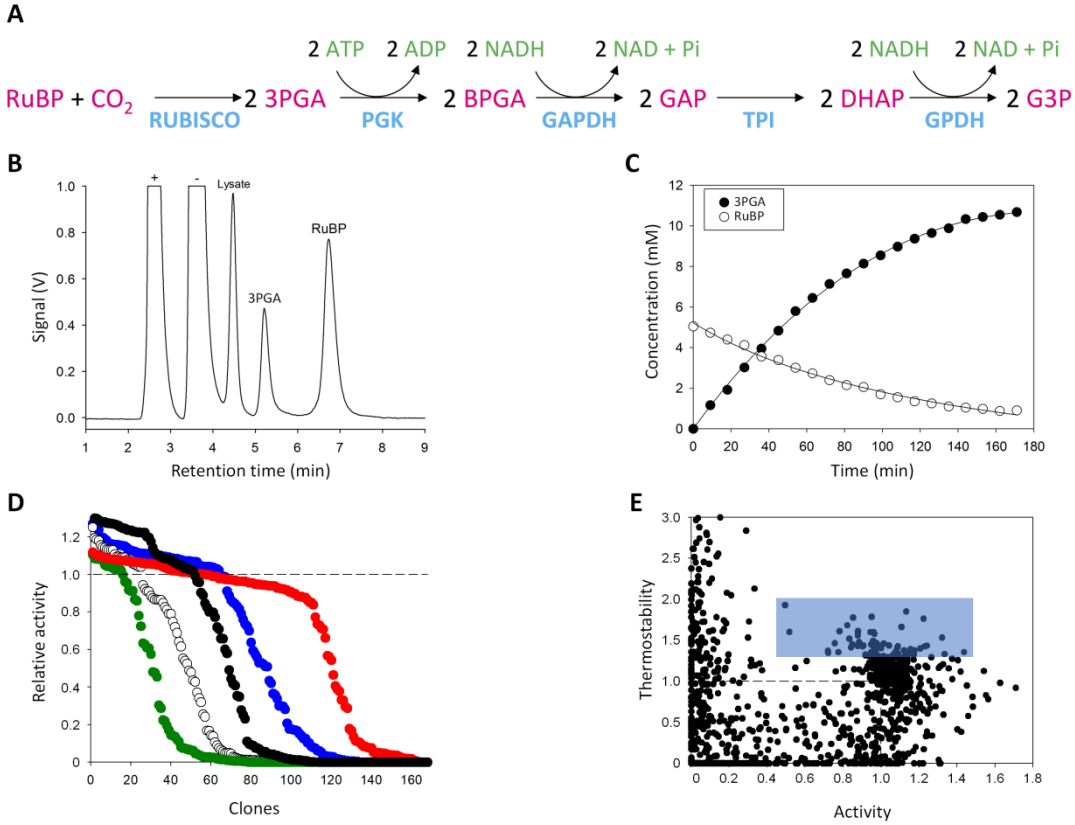


Figure 4.1.1 Dual HTS assay. (A) NADH-linked enzymatic assay of Rubisco activity. The reaction monitors the rate 3PGA production from the carboxylation of Ribulose-1,5-bisphosphate (RuBP) by Rubisco. PGK, phosphoglycerokinase; BPGA, 1,3-bisphosphoglycerate; GAPDH, glyceraldehyde-3P dehydrogenase; GAP, glyceraldehyde-3P; TPI, triose-P isomerase; DHAP, dihydroxyacetone-P; GPDH, glycerol-3P dehydrogenase; G3P, glycerol-3P. **(B)** HPLC-ELSD chromatogram for the detection of 3PGA production/RuBP depletion by Rubisco (+ and - represent cationic and anionic compounds in the sample, respectively) and **(C)**, the variation in the concentration of these two compounds during the reaction as measured by HPLC-ELSD. **(D)** Fitness landscapes for RubRr mutant libraries that are adjusted by varying the MnCl2 concentrations (in the case of Taq libraries) or by the amount of the DNA template (for mutazyme libraries). The relative RubRr activity of the clones is plotted in descending order and the dashed line shows the activity of the parental type in the assay: red circles, library I (mutazyme, 750 ng DNA template); blue circles, library II (Taq, 0.02 mM $MnCl_2$); black circles, library III (mutazyme, 100 ng DNA template); white circles, library IV (Taq, 0.03 mM $MnCl_2$); green circles, library V (Taq, 0.05 mM $MnCl_2$). **(E)** Directed evolution landscape for thermostability and activity relative to the parental RubRr enzyme (dashed lines). The mutants selected for further re-screening are

contained in the blue rectangle (using thresholds of 0.5- and 1.3-fold over RubRr activity and thermostability, respectively).

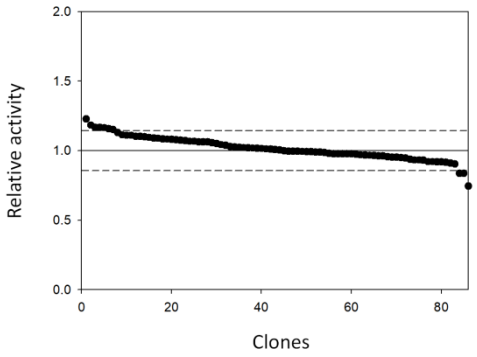


Figure 4.1.2 Validation of the spectrophotometric assay. Coefficient of variation (CV) for the spectrophotometric assay. The landscape corresponds to 88 independent clones containing RubRr parental type, grown in microtiter format, lysated and measured with the NADH depletion assay. The activity of the clones is plotted in descending order, solid line indicates the average value of the whole plate and dashed lines indicate the CV of the assay.

The *in vitro* evolutionary platform was firstly tested by constructing a small set of RubRr mutant libraries with different mutational loads and biases, **Figure 4.1.1D**. Mutational frequencies ranging from 12 to 73% of inactive clones were observed depending on the mutagenic PCR conditions applied. Several clones were randomly selected and sequenced to verify that the HTS assay and the mutational rates were adjusted correctly. The host-independent HTS assay and the expression system in microtiter format were used to evolve RubRr via different strategies (adaptive evolution, focused evolution, neutral genetic drift) in order to improve thermostability, **Figure 4.1.3A**. First, we determined the optimum temperature at which to stress RubRr mutant libraries. The T_{50} (the temperature where the enzyme lost 50% of its initial activity, after 10 minutes of incubation) was estimated in 64ºC from fresh lysate preparations and it was the main selecting force used during screening, **Figure 4.1.1E**.

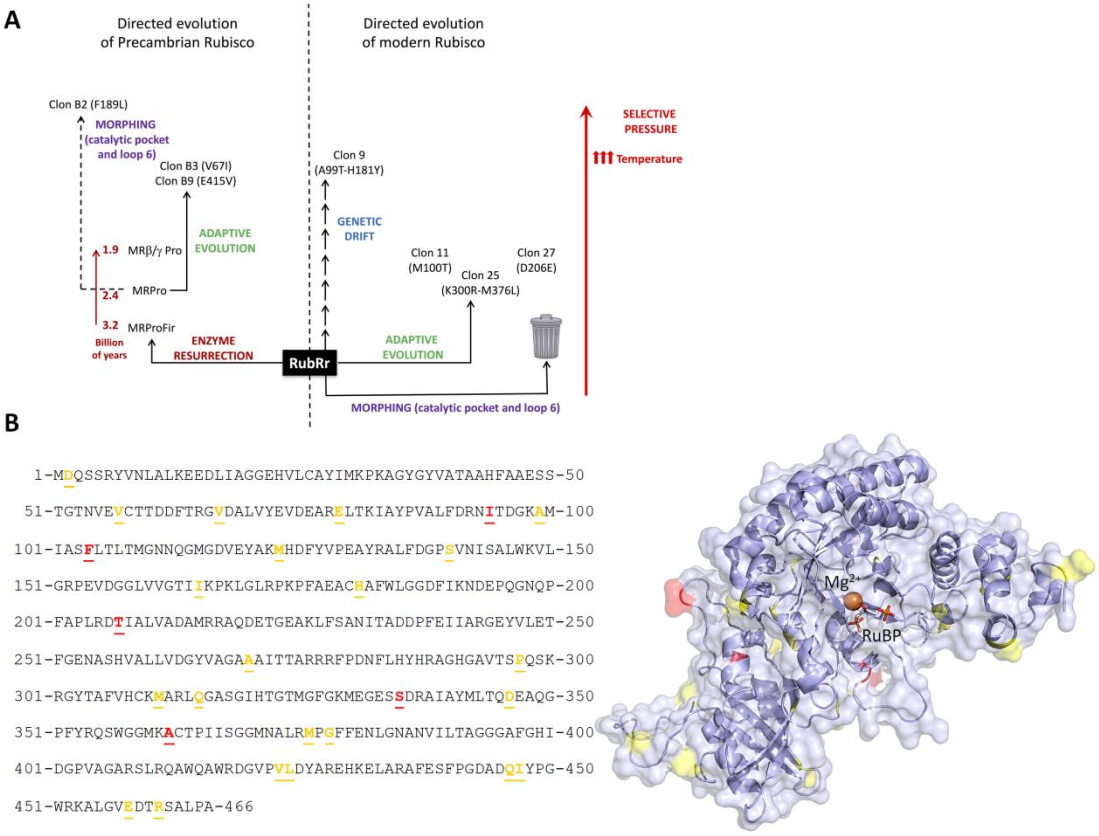


Figure 4.1.3 Directed -*in vitro*- evolution campaign. **(A)** General overview of the directed evolution campaigns carried out on Precambrian and modern (RubRr) form II Rubisco. **(B)** Neutral mutations identified in the amino acid sequence of RubRr (left) and mapped to its crystal structure (right). Mutations identified once are in yellow and those found in several clones are in red.

From this first adaptive evolution campaign, we identified three variants with improved stabilities (clone 11, clone 25 and clone 27, harboring M100T, K300R-M376L and D206E mutations, respectively), **Tables 4.1.1, 4.1.2**; **Figures 4.1.3A, 4.1.4, 4.1.5**.

Table 4.1.1 Biochemical characterization of extant and ancestral evolved Rubiscos.

Rubisco	*In vitro* Evolution strategy	T_m (°C)	[Rubisco] (% tsp)*	$S_{C/O}$ (mol.mol^{-1})	K_C (μM)	k_{cat}^C (s^{-1})	k_{cat}^C / K_C (mM.s^{-1})	K_m^{RuBP}(μM)
RubRr	Modern parent	70.7	28.1 ± 2.6	12.3 ± 0.2	105.2 ± 4.6	13.1 ± 1.1	125.2 ± 16	4.6 ± 1.5
Clone 9	Genetic drift	73.2	23.6 ± 1.3	11.6 ± 0.3	99.6 ± 1.1	12.9 ± 0.2	129.5 ± 2.9	5.4 ± 1.4
Clone 11	Adaptive evolution	74.5	31.0 ± 1.7	11.7 ± 0.2	318.7 ± 40.1	10.6 ± 0.9	33.6 ± 0.3	6.9 ± 0.9
Clone 25	Adaptive evolution	77.4	28.2 ± 1.2	12.7 ± 0.3	178.0 ± 4.1	8.1 ± 0.5	45.2 ± 1.9	45.9 ± 1.5
Clone 27	Adaptive evolution	74.2	34.6 ± 4.9	12.1 ± 0.2	111.5 ± 2.2	14.1 ± 1.1	126.5 ± 8.3	8.7 ± 0.8
MRPro	Ancestral parent	69.0	45.0 ± 4.7	6.5 ± 0.2	97.7 ± 3.2	4.0 ± 0.2	40.7 ± 1.1	158.3 ± 16.2
Clone B2	MORPHING	75.0	44.9 ± 6.9	6.9 ± 0.1	198.2 ± 15.4	1.4 ± 0.1	7.1 ± 0.7	375.1 ± 42.9
Clone B3	Adaptive evolution	70.0	41.4 ± 2.6	6.2 ± 0.1	105.1 ± 6.3	3.8 ± 0.5	36.3 ± 2.9	164.4 ± 13.8
Clone B9	Adaptive evolution	n.d.	56.5 ± 3.3	7.7 ± 0.6	94.9 ± 2.7	3.5 ± 0.1	36.6 ± 2.5	206.7 ± 14.8

*tsp, total soluble protein

Table 4.1.2 Structural information of mutations introduced in extant and ancestral Rubiscos.

Rubisco	Mutation	Secondary structure motif	Relative position	Distance to catalytic pocket (Å)
Clone 9	A99T	Loop (bridge)	Surface, N-terminal	20.5
	H181Y	α helix	Surface	10.2
Clone 11	M100T	Loop	N-terminal, interaction between both monomers	18.8
Clone 25	K300R	Loop	Surface, interaction of both monomers	26.9
	M376L	α helix	Surface, in the access site of the substrate	16.6
Clone 27	D206E	α helix	Surface	13.9
Clone B2	F189L	β sheet	Catalytic pocket	4.5
Clone B3	V67I	Loop	Surface, N-terminal	40.4
Clone B9	E415V	α helix	Surface	20.8

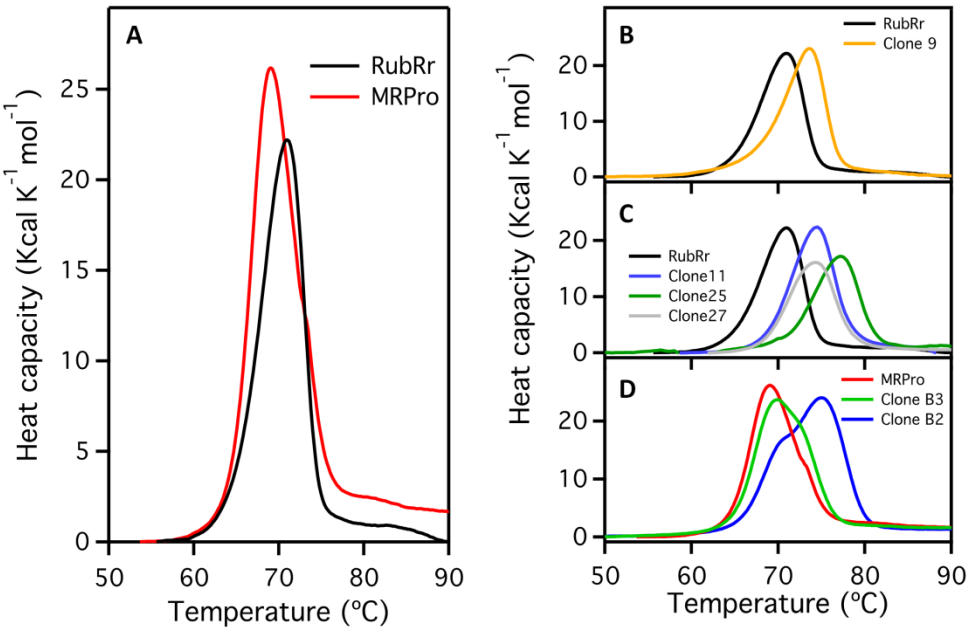


Figure 4.1.4 Differential scanning calorimetric (DSC) unfolding curves. (A) RubRr and MRPro. **(B)** RubRr and clone 9 (from genetic drift campaign). **(C)** RubRr and clones 11, 25 and 27 (from adaptive evolution). **(D)** MRPro and clones B2 and B3.

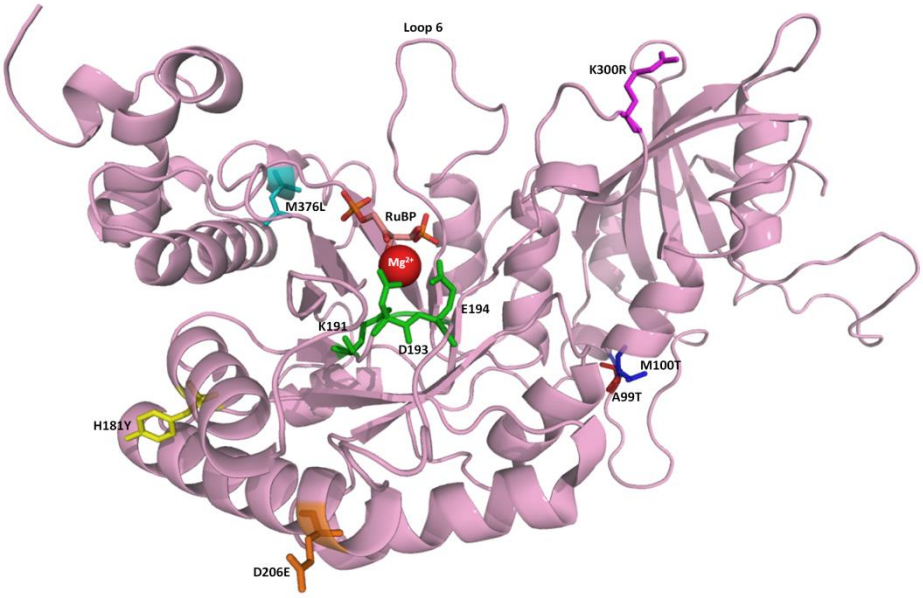


Figure 4.1.5 Molecular model of subunit RbcL1 (PDB code 9RUB) including the mutations introduced in modern RubRr. The Rubisco structure is shown as a pink cartoon while the RuBP substrate and the residues of the catalytic pocket are highlighted as orange and green sticks, respectively, and with Mg^{2+} as a red sphere. Numbering is relative to the RubRr sequence. The crystal structure of RubRr at a resolution 2.6 Å was used for modeling the mutants by Pymol (Schrodinger, LLC [http://www.pymol.org]).

It is worth noting that two of these four substitutions (M376L and D206E) were identified as ancestor mutations, as they were present in the ancestral nodes reconstructed from RubRr (see below, **Figure 4.1.6, 4.1.7**), which is in excellent agreement with the stabilization effect observed when inserting back-to-ancestor mutations (Risso et al. 2014; Bershtein et al. 2008).

We explored the versatility of the *in vitro* evolution platform to construct mutant libraries by neutral genetic drift, since it has been proved to be a valuable evolution strategy to enhance protein stability (Bloom et al. 2007). In this case, by accumulating neutral mutations that maintain wildtype function, functional polymorphic populations are rising up while purging detrimental mutations, a process also known as purifying selection (Gupta and Tawfik, 2008). As a rule of thumb, we established a cut-off >75% of the wildtype´s activity in the mutational landscape so that around 200 functional neutral clones per round of evolution were selected, pooled, and subjected to further random mutagenesis. After screening over 6,000 clones in seven generations of neutral drift, we picked randomly 20 neutral variants from the last generation, which were sequenced and analyzed, **Figure 4.1.3B**. The average number of neutral mutations found per clone -excluding silent

mutations- was 1.65 which agrees with the mutational windows described in former studies (Kaltenbach and Tokuriki, 2014). Neutral mutations were mapped onto the crystal structure of RubRr. As expected most of them were found at the surface of the protein, far from the catalytic motives with the exception of I165V (in the catalytic pocket) and M311V, Q315H and S335G (located in conserved flexible loop 6 of the C terminal), **Figure 4.1.3B**. Among the neutral mutants analyzed, the clone 9 (A99T-H181Y) improved both kinetic and thermodynamic stabilities, **Tables 4.1.1, 4.1.2**; **Figures 4.1.3A, 4.1.4, 4.1.5** (see biochemical characterization below).

4.1.2 Directed -*in vitro*- evolution of Precambrian Rubisco

The prediction of ancestral enzymes by phylogenetic algorithms is a plausible approach to recreate version phenotypes that were key in the metabolism of primitive cells (Shih et al. 2016; Perez-Jimenez, R. 2011; Gaucher et al. 2008). Very recently, it has been suggested that the synthetic courtship between ancestral enzyme resurrection and directed evolution could yield promising blueprints for adaptation to different environments and new functions (Alcalde, 2017b; Risso and Sanchez-Ruiz, 2017). Given that there are no studies on ancestral Precambrian enzymes evolved in the laboratory, and that Rubisco is among the most ancient enzymes known in earth history, we planned the directed evolution of predicted Precambrian Rubisco. We first carried out the reconstruction of several ancestral nodes of Rubisco form II: the most recent common ancestor of β and γ proteobacteria (MRβ/γPro), the most recent common ancestor of all the proteobacteria (MRPro) and the most recent common ancestor of proteobacteria and the cyanobacteria firmicutes (MRProFir), which dated back from 1.9 to 3.2 Gyr ago, **Figures 4.1.3A, 4.1.6, 4.1.7**.

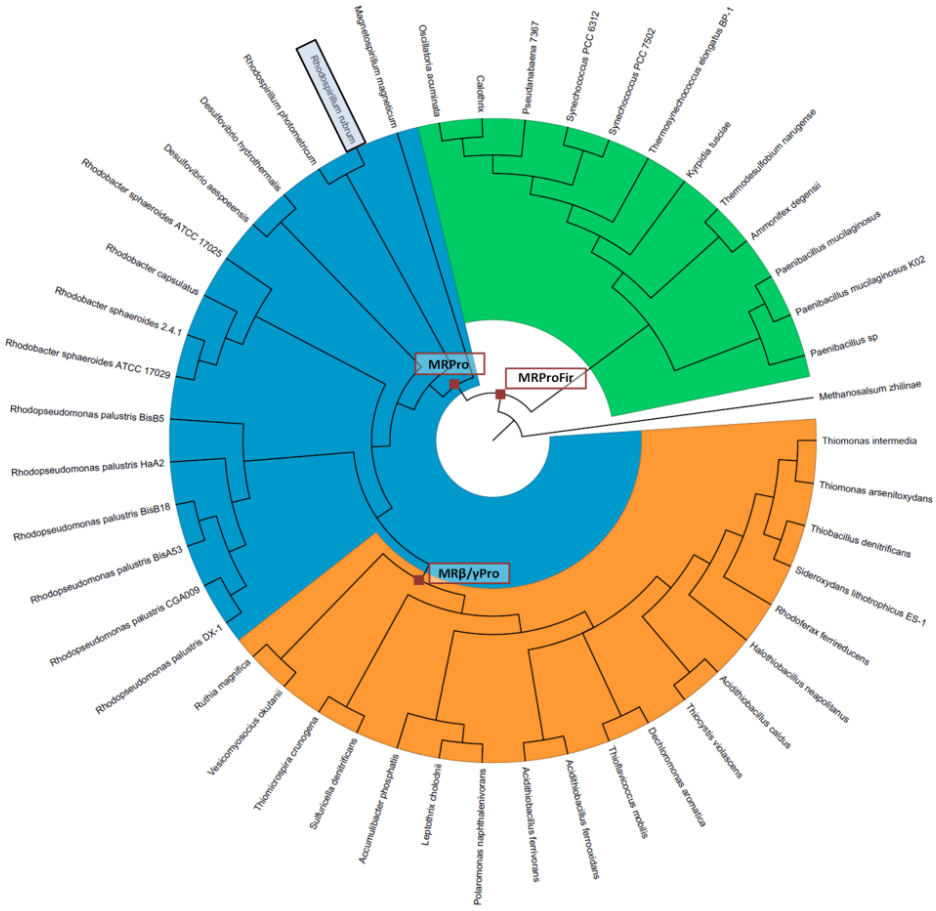
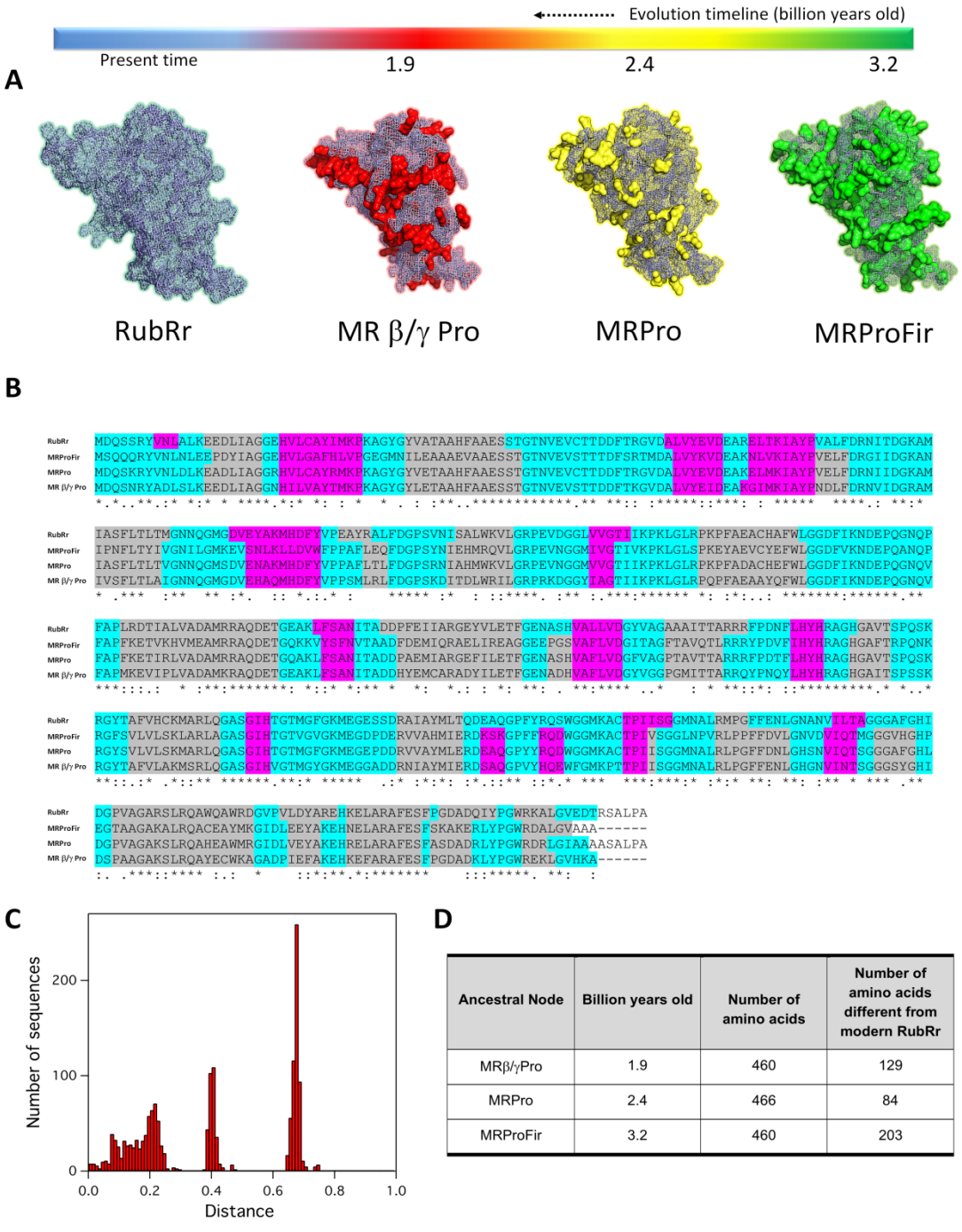


Figure 4.1.6 Phylogenetic tree generated from the amino acid sequences of 47 different form II Rubisco sequences. The colored areas represent the different bacterial phyla and classes as follows: Firmicutes (Green); α and δ proteobacteria (Blue); γ and β proteobacteria (orange). The brown squares highlight the nodes whose sequences were selected for resurrection and the reference *R. rubrum* Rubisco (RubRr) sequence is framed.

These three ancestral nodes were successfully resurrected (*i.e.* functionally expressed in *E. coli*). The preliminary biochemical characterization of ancestral nodes revealed MRPro (2.4 Gyr) as the most promising departure point for directed evolution due to its higher activity levels, **Figure 4.1.8A**. In order to compare the evolution *in vitro* of ancestral and modern Rubisco, we constructed independent mutant libraries using MRPro and modern RubRr as templates whereby identical mutational rates of 3-7 mutations per kilobase (kb) were applied -confirmed by DNA sequencing a random sample of clones-, a number that is duplicated in mature Rubisco in its dimeric functional complex L_2. Interestingly, the resulting evolution landscapes were broadly different, with a 76% of inactive clones for the extant rubisco vs. only 26% of unfolded variants for the ancestral counterpart, **Figure 4.1.8B**.



D

Ancestral Node	Billion years old	Number of amino acids	Number of amino acids different from modern RubRr
MRβ/γPro	1.9	460	129
MRPro	2.4	466	84
MRProFir	3.2	460	203

Figure 4.1.7 Comparison between modern RubRr and the ancestral -resurrected- Rubisco. **(A)** Amino acid differences (in surface mode) in MRProFir (green), MRPro (yellow) and MRβ/γPro (red) relative to RubRr. **(B)** Multiple sequence alignment of RubRr and ancestral Rubiscos. Expected motifs are highlighted in colors. (pink, beta sheets; grey, alpha helices; blue, loops). The asterisk indicates positions with a fully conserved residue; colon indicates conservation between groups of strongly similar properties; period indicates conservation between groups of weakly similar properties. **(C)** Frequency distribution of the pairwise sequence distances for form II Rubisco sequences obtained from a database search using the RubRr as the query. The sequence distances are represented on the X axis and the number of sequences is reflected on the Y axis. **(D)** Number of amino acids different from RubRr and billion years old between the ancestral nodes. The multiple sequence alignment was performed with Clustal Omega V1.2.4 available at https://www.ebi.ac.uk/Tools/msa/clustalo/. The structures were modeled using the PDB code 9RUB and the Phyre2 server (Protein Homology/analogY Recognition Engine V 2.0: Kelley and Sternberg, 2009) available at www.sbg.bio.ic.ac.uk/phyre2.

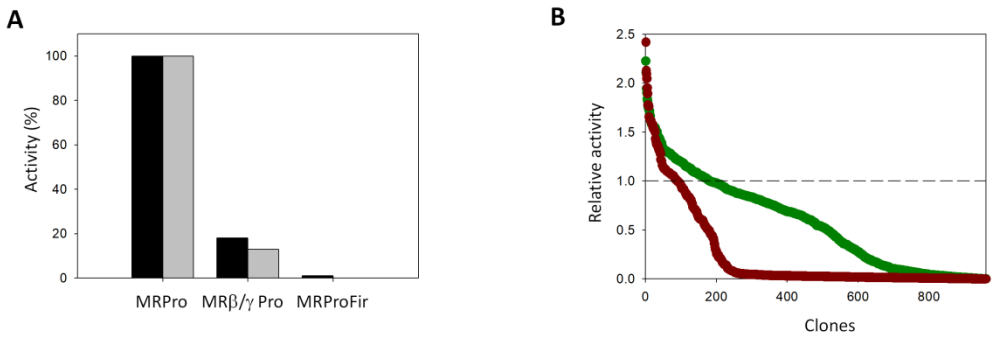


Figure 4.1.8 Activity of ancestral nodes and evolution landscapes. (A) Activity of the ancestral Rubisco nodes relative to MRPro (in %) measured in a microplate (grey) and flask (black) format. **(B)** Evolution landscapes for RubRr (red circles) and MRPro (green circles) adaptive mutant libraries. The activity of the clones is plotted in descending order and the dashed line shows that of the parental type.

This result highlights the striking marked tolerance of MRPro to the introduction of random mutations, which supports the hypothesis of the potential evolvability of ancestral proteins to obtain more beneficial phenotypes (Risso and Sanchez-Ruiz, 2017). Then, MRPro was included within the directed evolution program for stability returning with the thermostable clone B3 (with V67I mutation), **Tables 4.1.1, 4.1.2**; **Figure 4.1.3A, 4.1.4, 4.1.9**. To compare the evolvability of Precambrian and modern Rubiscos, MRPro and RubRr were jointly subjected to structure-guided evolution by MORPHING, a focused random mutagenesis method developed in our laboratory for directed evolution in yeast that was adapted for this occasion to bacteria (Gonzalez-Perez et al. 2014). By MORPHING, we targeted short structural motives to random mutation without altering the remaining regions of the protein. Given that mutating active site residues may enhances stability albeit at the cost of activity, after computational analysis we concentrated MORPHING onto two highly sensitive regions: i) the core catalytic pocket -particularly in a sequence block of 55 residues (K148-L204, according to RubRr numbering) that contains several amino acids directly involved in catalysis (*i.e.* K166, K168, K191, D193, E194), and ii) in the loop 6 -a twelve amino acid segment G323-D336 that shows different conformational states upon RuBP binding, **Figure 4.1.10** (Parry et al. 2003; Duff et al. 2000; Taylor and Andersson, 1997; Schneider et al. 1989).

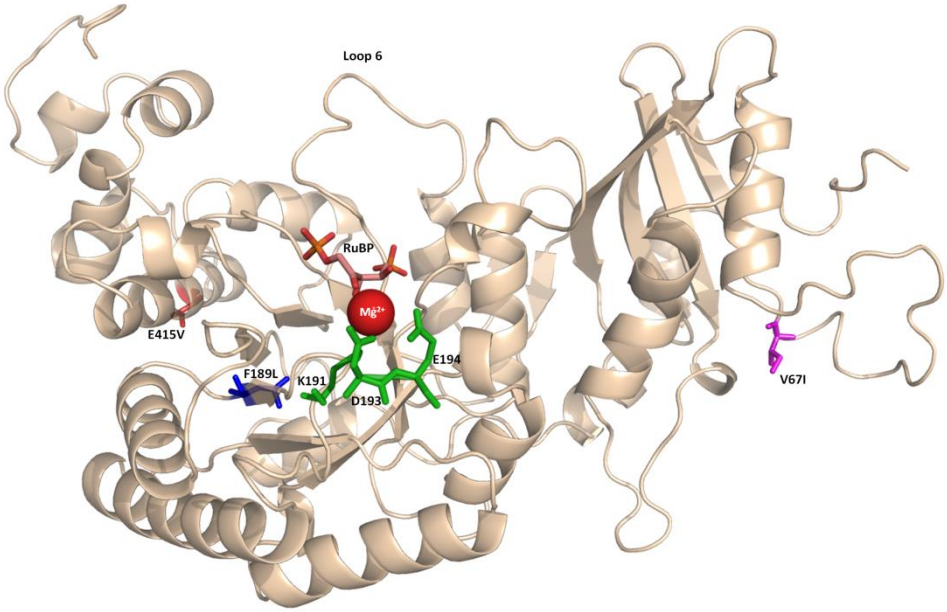


Figure 4.1.9 Molecular model of subunit RbcL1 including the mutations introduced in ancestral MRPro. The Rubisco structure is shown as a light brown cartoon while the RuBP substrate and the residues of the catalytic pocket are highlighted as orange and green sticks, respectively, and with Mg^{2+} as a red sphere. Highlighted are the positions of mutations F189L (clone B2), V67I (clone B3) and E415V (clone B9). The model was made using the Phyre2 server (Protein Homology/analogY Recognition Engine V 2.0: Kelley and Sternberg, 2009) available at www.sbg.bio.ic.ac.uk/phyre2. The outcome model was used for modeling the mutants by Pymol (Schrodinger, LLC [http://www.pymol.org]).

While this mutational strategy was too aggressive for extant RubRr (with a great percentage of inactive clones and none selected winners), functional mutant libraries of MRPro were obtained at both targeted regions, with clone B2 (with F189L substitution) as the most stable mutant of the focused evolution experiment, **Tables 4.1.1, 4.1.2**; **Figures 4.1.3A, 4.1.4, 4.1.9**.

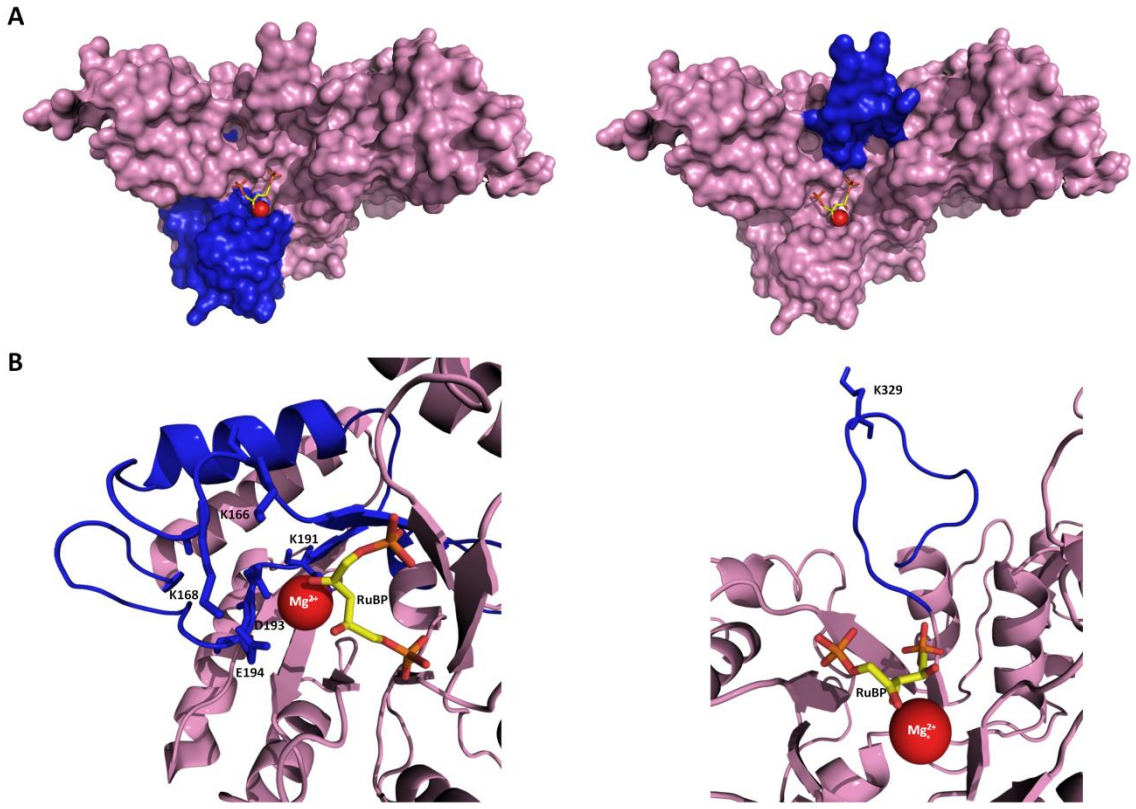


Figure 4.1.10 MORPHING in RubRr. (A) Surface of the subunit RbcL1 with the regions selected for MORPHING highlighted in blue and the non-mutated areas in pink: left, catalytic pocket region; right, loop 6 region. **(B)** Detail of the secondary motifs that form the catalytic pocket (Left) and loop 6 (Right). Some of the conserved residues involved in catalysis (catalytic pocket) and in the closure over the substrate (loop6) are labeled, and highlighted as sticks.). The model was made using the PDB code 9RUB.

4.1.3 Biochemical characterization

The ancestral node MRPro, the modern RubRr and their corresponding offspring were fully characterized in terms of folding, thermodynamic stability and kinetic parameters. Rubisco expression was quantified by ^{14}C-CABP (2-carboxyarabinitol-1,5-bisphosphate) inhibitor binding and confirmed by SDS-PAGE, **Table 4.1.1**; **Figures 4.1.11, 4.1.12**. Strikingly, MRPro and its evolved mutants (clones B2, B3 and B9) showed a strong solubility ranging from 45% (in MRPro) up to 56% (in clone B9) of the total soluble protein (tsp) *vs.* ~30% of modern counterparts, **Table 4.1.1**; **Figure 4.1.11**. This easiness of Precambrian Rubiscos in functional expression and assembly in a fully dimeric functional state within a modern microorganism may be related to its predicted ancestral origin (2.4 Gyr ago), a period in which primitive cells relied on only a small set of promiscuous enzymes such as enough expression levels had to be guaranteed to fulfill most metabolic tasks. All stabilizing mutations were found at the surface of the protein, far from

catalytically relevant regions with the exception of F189L of clon B2, **Table 4.1.2**; **Figures 4.1.5, 4.1.9.**

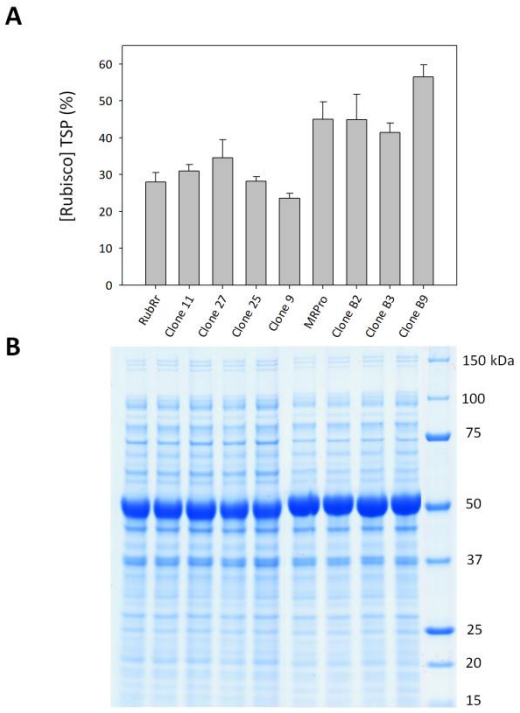


Figure 4.1.11 Expression of RubRr, ancestral Rubisco and their mutant offspring in *E. coli*. **(A)** Rubisco content as a percentage of the total soluble *E. coli* protein (TSP, %) determined by 14C-CABP binding and confirmed by **(B)** SDS-PAGE. Lane 1, RubRr; Lane 2, clone 11; Lane 3, clone 27; Lane 4, clone 25; Lane 5, clone 9; Lane 6, MRPro; Lane 7, clone B2; Lane 8, clone B3; Lane 9, clone B9; Lane 10, Molecular weight (MW) protein standard. The Rubisco RbcL subunit corresponds to the prominent 50 kDa band.

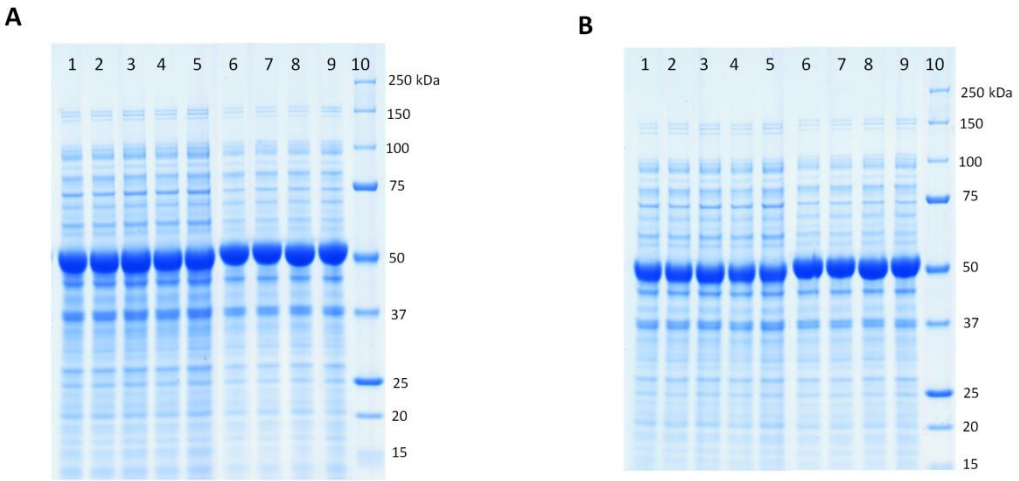


Figure 4.1.12 SDS PAGE analysis of Rubisco production in *E. coli*. a) Total cellular and **b)** soluble cellular protein separated by SDS-PAGE highlighting the expression and solubility of RubRr, MRPro and their derived mutants in *E coli*. Lane 1, RubRr; Lane 2, clone 11; Lane 3, clone 27; Lane 4, clone 25; Lane 5, clone 9; Lane 6, MRPro; Lane 7, clone B2; Lane 8, clone B3; Lane 9, clone B9; Lane 10, MW protein standard. The Rubisco RbcL subunit corresponds to the 50kDa band. Refer to Fig 2a for the origins and mutations of each clone.

The detailed analysis of thermodynamic stabilities of selected thermostable variants (clones 11, 25, 27 and 9 from RubRr evolution; clones B2 and B3 from MRPro evolution) confirmed the improvements observed during the screenings of the different *in vitro* evolution experiments. When the adaptive evolution experiment was performed on modern RuBRr, clone 25 outperformed parental type by ~7ºC in the T_m (determined by DSC-differential scanning calorimetry) followed by clones 27 and 11, **Table 4.1.1**; **Figure 4.1.4**. The clone 9 from the neutral drift campaign showed a 2.5ºC improvement in thermodynamic stability. Clone B2, on which we targeted the catalytic site by focused evolution, displayed a significant enhancement of ~6ºC in the T_m, albeit at the cost of kinetics (as it happened, to a lesser extent, with some of the other variants). In this regard, the inherent biochemical trade-off between the carboxylation efficiency k_{cat}^{C}/ K_C (that comes from dividing the rates of CO_2-fixation (k_{cat}^{C}) by the affinity for CO_2 under no O_2 atmosphere (K_C) at 25ºC) and the specificity for CO_2 over O_2 ($S_{C/O}$) became apparent, something that could be expected if we bear in mind that these properties were not addressed in the screen, **Table 4.1.1 and Figure 3.3.4 (from Materials and Methods section)**. Ancestral variants showed a poor $S_{C/O}$ when compared with modern counterparts, which is consistent with the low selective pressure that Precambrian Rubisco found on primitive earth (in terms of atmospheric disproportion between CO_2 and O_2), right before the first biogeochemical dawn of O_2 in atmosphere (estimated 2.5 Gyr ago) (Sessions et al. 2009). Along these lines, the low carboxylation efficiency shown by ancestral Rubiscos agrees well with recent resurrected Rubiscos type I from the Mesoproterozoic era (> 1 Gyr ago) addressing them as reliable proxies of true ancestral sequences (Shih et al. 2016).

4.2 Chapter II: Ancestral Resurrection and Directed Evolution of Fungal Paleozoic Laccases

Ancestral sequence reconstruction and resurrection is a useful tool for protein engineering whose alliance with directed evolution has hardly been explored. In this study we have resurrected several ancestral nodes of fungal laccases dated back 500-250 million years. Unlike modern laccase, the Paleozoic laccases were readily secreted by yeast, with similar kinetic parameters, broader stability as well as modified pH activity profiles, a possible molecular memory of their ancestral origin. The resurrected laccase, LacAnc100, carrying 136 ancestral mutations, was subjected to directed evolution to improve the oxidation rates against 1,3 cyclopetanedione, a β–diketone initiator commonly used in vinyl polymerization reactions.

This chapter is based on the article under preparation: Gomez-Fernandez B.J., Risso V.A., Rueda A., Sanchez-Ruiz J.M. and Alcalde M. (2019). Directed evolution of an ancestral laccase.

4.2.1 Reconstruction and resurrection of ancestral laccases

Using as a query sequence the high-redox potential laccase from basidiomycete PM1 (Coll et al. 1993), we reconstructed three ancestral laccase nodes from the Paleozoic: LacAnc95, LacAnc98 and LacAnc100, **Figure 4.2.1**. The TimeTree of Life was employed to locate the ancestral nodes at the phylogenetic level, dating them back 500-252 million years, within the appearance of the common ancestor of Basidiomycota, in the early Cambrian, and the common ancestors of Agaricomycotina and Agaricomycetes, **Figure 4.2.2** (Hedges and Kumar 2009). Indeed, this period was extremely rich in new species due to an evolutionary burst that matched with the warm up of earth (Sperling et al. 2013) so that new biochemical traits arose. The three ancestral nodes were cloned in *S. cerevisiae*, the most common host used for directed evolution of fungal laccases (Mate and Alcalde, 2015). To foster secretion, laboratory evolved and/or chimeric versions of the α-factor prepro-leader from *S. cerevisiae* (α^{PM1}, α^{PcL}, pre-α-Prokiller), were attached to each node and evaluated (see Materials and Methods for details) (Mateljak et al. 2017; Viña-Gonzalez et al. 2015; Camarero et al. 2012; Mate et al. 2010). In order to optimize culture conditions, all constructs were screened in microtiter fermentations of the supernatants (*i.e.* cultures in 96 well plates). Both LacAnc98 and LacAnc100 fusions attached to α^{PcL} displayed the highest activity levels from supernatants, with ethanol supply and at 30°C of incubation

temperature, produced in microtiter plates -*i.e.* easily detectable for a directed evolution campaign-, whereas LaAnc95 did not express irrespective of the signal peptide or culture conditions.

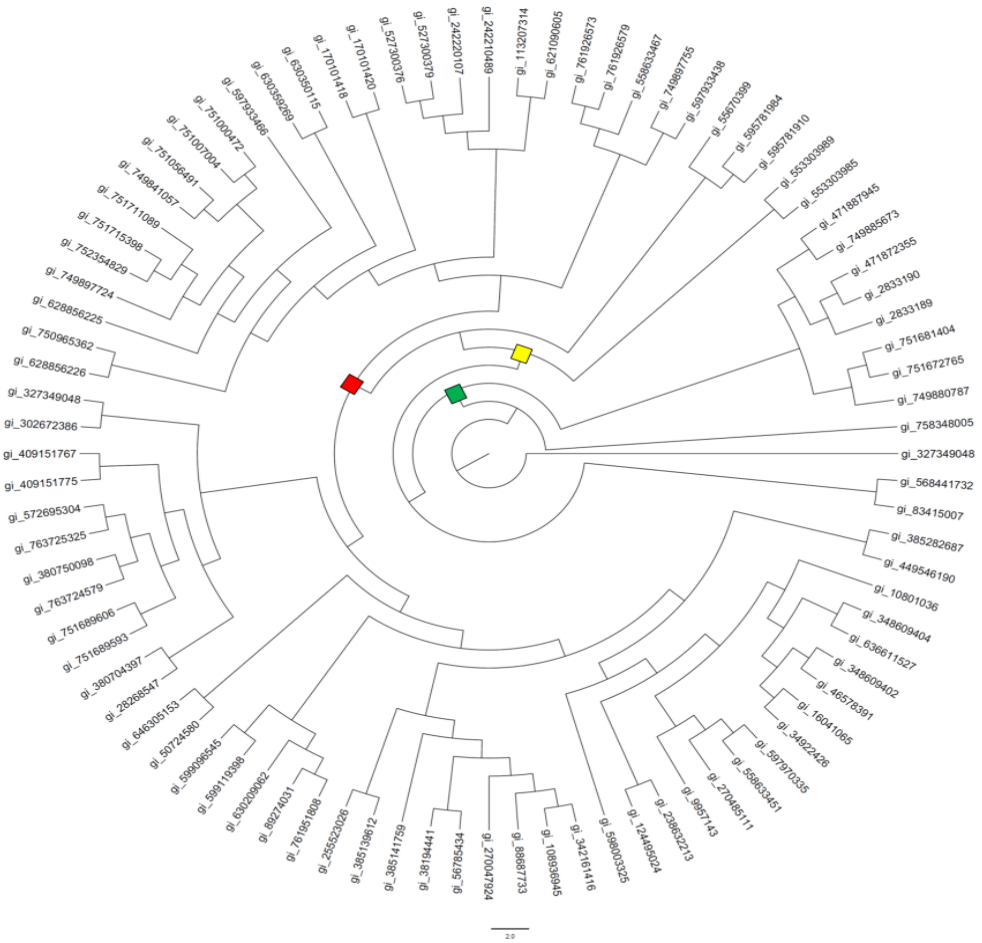
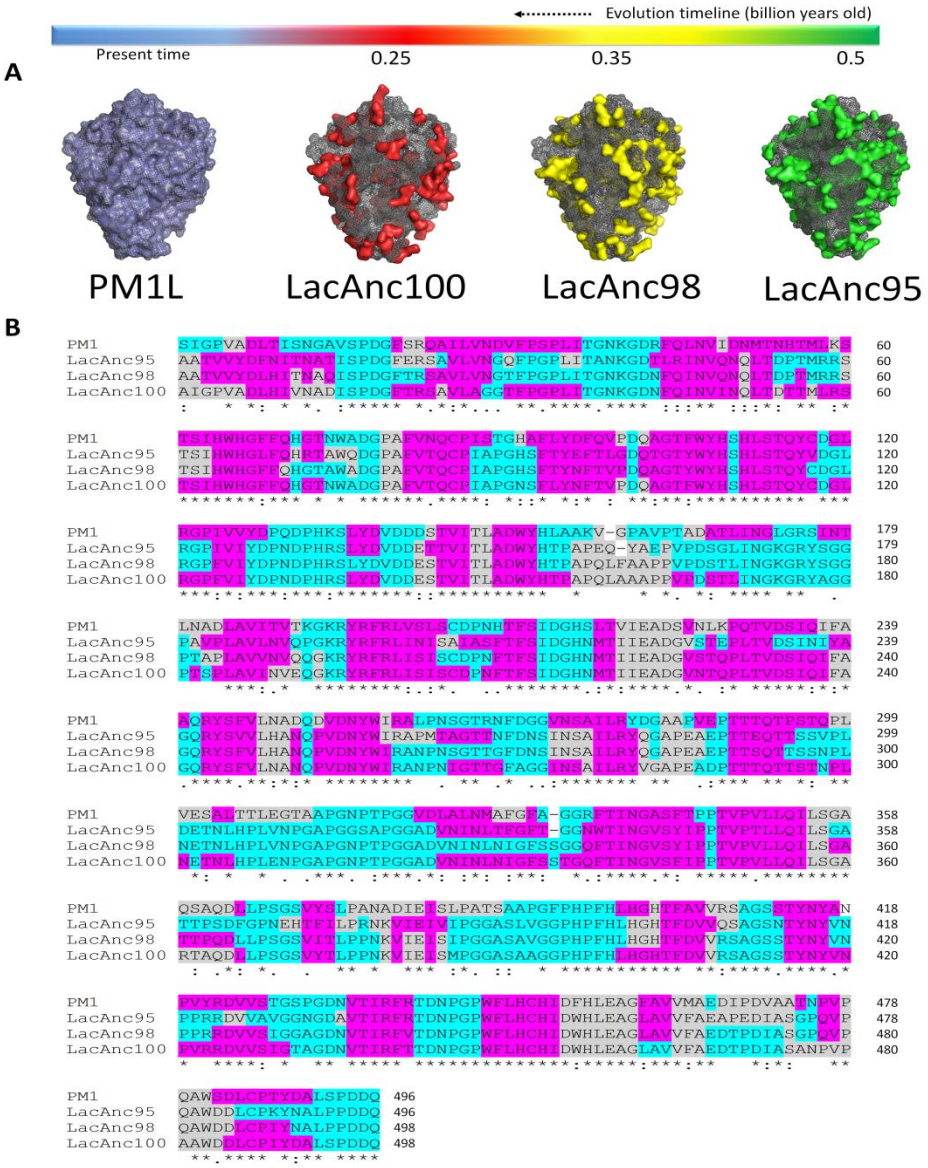


Figure 4.2.1 Phylogenetic tree generated from the amino acid sequences of 87 different laccases sequences. The nodes whose sequences were selected for resurrection are depicted as squares. Green square, LacAnc95; Yellow square, LacAnc98; Red square, LacAnc100.



C

Ancestral Nodes	Billion years old	Number of amino acids	Number of amino acids different from modern laccase
LacAnc95	~500	496	188
LacAnc98	~350-300	498	155
LacAnc100	~250	498	134

Figure 4.2.2 Comparison between modern laccase and the ancestral -resurrected- laccases. **(A)** Amino acid differences (in surface mode) in LacAnc95 (green), LacAnc98 (yellow) and LacAnc100 (red) relative to PM1L. **(B)** Multiple sequence alignment of PM1L and ancestral laccases. Expected motifs are highlighted in colors. (pink, beta sheets; grey, alpha helices; blue, loops). The asterisk indicates positions with a fully conserved residue; colon indicates conservation between groups of strongly similar properties; period indicates conservation between groups of weakly similar properties. **(C)** Number of amino acids different from PM1L and billion years old between the ancestral nodes. The multiple sequence alignment was performed with Clustal Omega V1.2.4 available at https://www.ebi.ac.uk/Tools/msa/clustalo/. The structures were modeled using the PDB code 5ANH and the Phyre2 server (Protein Homology/analogY Recognition Engine V 2.0: Kelley and Sternberg, 2009) available at www.sbg.bio.ic.ac.uk/phyre2

The LacAnc98 and LacAnc100 were produced, purified to homogeneity and biochemically characterized. Unfortunately, we couldn't compare the ancestral nodes with the native basidiomycete PM1 laccase produced by *S. cerevisiae* due to its poor secretion (0.035 ABTS-Units/L of supernatant) (Mate et al. 2013). Instead, we decided to use the evolved OB-1 laccase for benchmarking with the ancestral nodes, which is the product of eight generations of directed evolution from PM1 laccase, carrying mutations V[α10]D-N[α23]K-A[α87]T-V162A-H208Y-S224G-A239P-D281E-S426N-A461T (the mutations lying in the evolved α-factor prepro-leaders are underlined) that conferred improved secretion and activity in *S. cerevisiae* (Mate et al. 2010). The molecular mass of LacAnc98 and LacAnc100 determined by SDS-PAGE was ~60 kDa with a glycosylation degree of 15%, similarly to the evolved OB-1, (Mate et al. 2010), **Figure 4.2.3**. To assess correctly expression levels, OB-1 had to be fused to α^{PcL} (the same prepro-leader attached to the Paelozoic nodes). Ancestral fusions showed a 10-fold higher secretion than OB-1. This result is striking if we consider the general difficulties found when expressing native fungal laccases heterologously, which typically need of exhaustive directed evolution campaigns to achieve reasonable secretion levels (Camarero et al. 2012; Mate et al. 2010; Bulter et al. 2003).

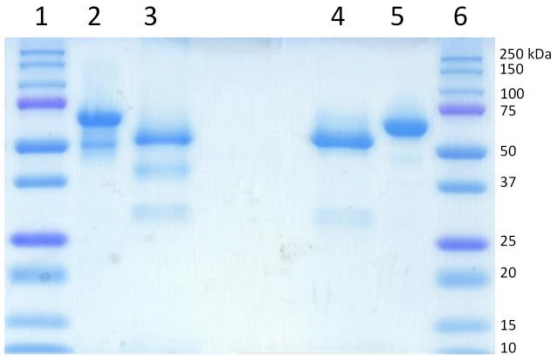


Figure 4.2.3 SDS-PAGE of purified LacAnc98 and LacAnc100. Lanes: 1, protein ladder; 2, LacAnc98; 3, deglycosylated LacAnc98; 4, deglycosylated LacAnc100; 5, LacAnc100; 6, protein ladder. The purified enzymes were desglycosylated using Peptide N-glycanase (PNGase F). Samples were resolved on 12% SDS-polyacrylamide gel and stained with ProtoBlue Safe.

Kinetic stability was assessed by measuring the half-life of inactivation ($t_{1/2}$) which was defined as the time required by the laccase to lose 50% of its residual activity after incubation at a given temperature. When stability was measured at 50°C and pH 3.0, both ancestral nodes outperformed the $t_{1/2}$ of the modern laccase by roughly 55 min, **Figure**

4.2.4A. By contrast, the $t_{1/2}$ at 70°C and pH 6.0 was 15 min higher for OB-1, which indicates a strong effect of the selected pH on the overall stability, **Figure 4.2.4B**.

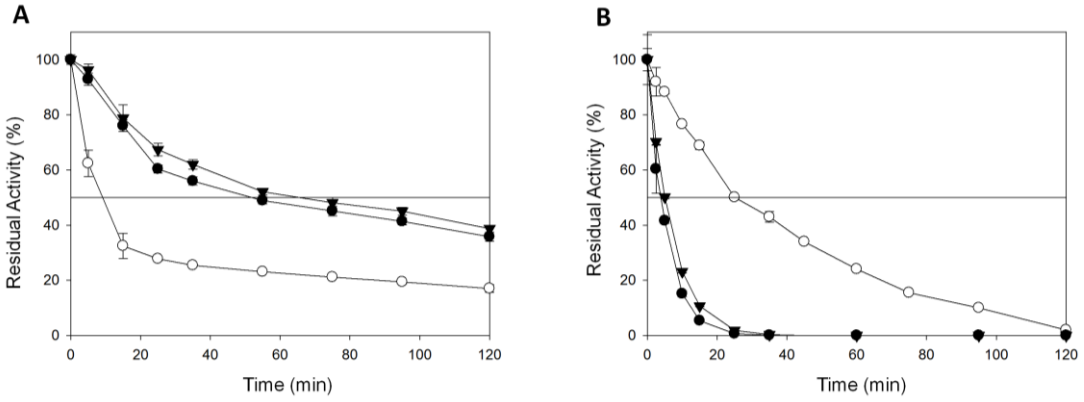


Figure 4.2.4 Half-live of LacAnc98, LacAnc100 and OB-1 variant in (A) pH 3.0 at 50°C and **(B)** pH 6.0 at 70°C. Solid line shows the residual activity at 50%. White circles, OB-1; Black circles, LacAnc98; Black triangles, LacAnc100. Each point, including the standard deviation, comes from three independent experiments.

In the light of these results, we studied pH stability at room temperature (in the pH range from 2.0 to 9.0) obtaining an overall hierarchy LacAnc100>LacAnc98>OB-1, **Figure 4.2.5A-C**. Specially, LacAnc100 kept good stability even at acidic pH (after long incubation periods of 144h at pH 2.0-3.0) where the residual activity of modern laccase was hardly measurable.

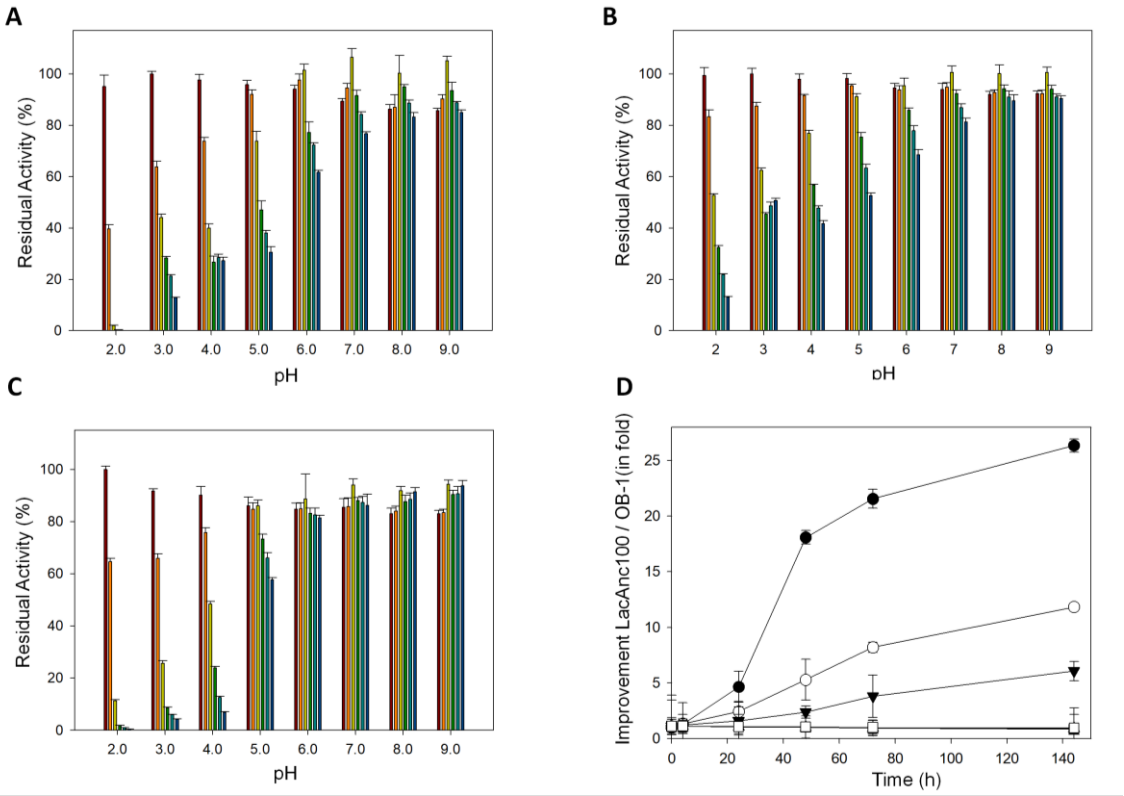


Figure 4.2.5 pH stability of **(A)** LacAnc98 **(B)** LacAnc100 and **(C)** OB-1 variant. Laccases were incubated for 0 (brown), 4 (orange), 24 (yellow), 48 (green), 72 (light blue) and 144 (blue) hours at different pHs values (2.0-9.0). Laccase activity was normalized to the highest activity value at time 0. Each point and standard deviation is from three independent measurements. **(D)** Difference in improvement (in fold) of residual activities between LacAnc100 and OB-1 after incubation at different pH. After the incubation of 0, 4, 24, 48, 72, 96 and 144 h at different pH values (2.0-7.0), residual activity was measured with 1 mM ABTS in 100 mM sodium phosphate/citrate buffer, pH 4.0. Residual activities of LacAnc100 were divided over residual activities of OB-1 variant at corresponding pH value to determine fold of improvement. Black circles, 2.0; White circles, 3.0; Black triangles 4.0; White triangles, 5.0; Black squares, 6.0; White squares, 7.0. Each point and the standard deviation come from three independent measurements.

As shown in **Figure 4.2.5D**, this improvement in the stability of LacAnc100 relative to OB-1 was enhanced by the gradual increase in acidity. The pH activity profile was measured for phenolic (2,6- dimethoxyphenol -DMP- and guaiacol) and non-phenolic ((2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)) -ABTS-) compounds as well as for potassium octacyanomolybdate- $K_4[Mo(CN)_8]$ (an inorganic transition metal complex), **Figure 4.2.6**. A deviation of LacAnc100 (and in a lesser extent for LacAnc98) towards more alkaline pH was observed for all substrates except for ABTS, with a shift in the optimum pH of activity, from 4.0 to 5.0, for the two phenolic substrates.

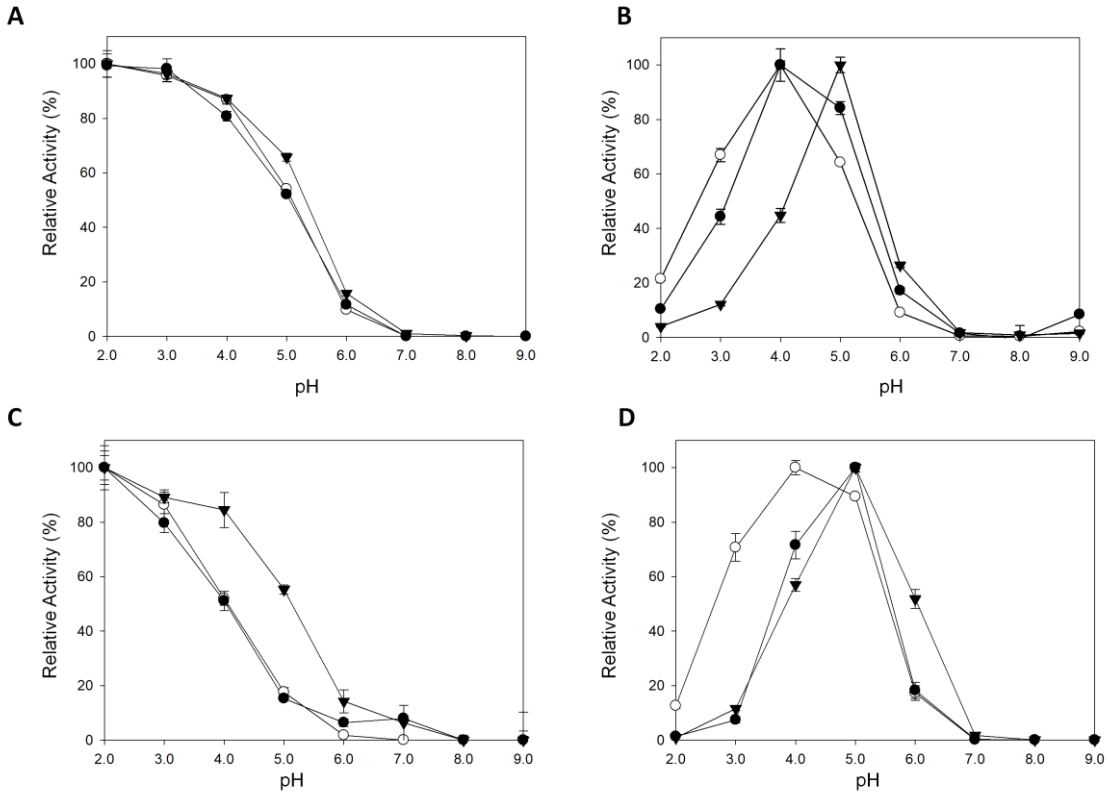


Figure 4.2.6 pH activity profile. White circles, OB-1; Black circles, LacAnc98; Black triangles, LacAnc100. Activities were measured in 100mM citric/phosphate/borate buffer at different pHs with 1mM ABTS **A)**, 5mM DMP **B)**, 5mM $K_4[Mo(CN)_8]$ **C)** or 50mM Guaiacol **D)** as the substrates. Laccase activity was normalized to the optimum activity value and each point, including the standard deviation, comes from three independent experiments.

Steady kinetic constants were measured with a panel of laccase substrates, **Table 4.2.1**. The catalytic efficiency of LacAnc98 was around one order of magnitude lower than that of OB-1 for all the substrates assayed. Conversely, LacAnc100 and OB-1 showed rather similar kinetic values, except for ABTS where the K_m for the ancestral laccase was 10-fold higher than in the extant one. It is important to note that the kinetic constants of native laccase PM1 homologously produced by the fungus are in the same order as those for OB-1 (for ABTS and DMP), which was intensively evolved to achieve high activity and secretion in yeast. This implies that the activity of native laccase PM1 must be several folds lower when the enzyme is expressed in *S. cerevisiae* than in the original fungus, as also happened for other laccases expressed in yeast (Bulter et al. 2003).

Table 4.2.1 Kinetic parameters of OB-1, PM1wt (from the fungus), LacAnc100 and LacAnc98. All reactions were performed by triplicate.

Substrate	Kinetic constant	OB-1	PM1wt (fungus)*	LacAnc100	LacAnc98
ABTS	K_m (mM)	0.007 ± 0.0006	0.0081 ± 0.0007	0.06 ± 0.001	0.0698 ± 0.0024
	k_{cat} (s^{-1})	690.0 ± 20.2	272 ± 7	1325.6 ± 0.06	468.44 ± 9.55
	k_{cat}/K_m (mM^{-1} s^{-1})	106073.8	33580	23054.2	6711.17
DMP	K_m (mM)	0.15 ± 0.01	0.12 ± 0.01	0.5 ± 0.02	0.4184 ± 0.212
	k_{cat} (s^{-1})	565.0 ± 13	153 ± 4	600.2 ± 10.3	238.39 ± 5.4
	k_{cat}/K_m (mM^{-1} s^{-1})	3842.6	1093	1217.3	569.76
Sinapic Acid	K_m (mM)	0.32 ± 0.02	0.048 ± 0.001	0.33 ± 0.07	0.227 ± 0.0425
	k_{cat} (s^{-1})	741.0 ± 14.1	45 ± 3	519.5 ± 55.2	129.20 ± 11.05
	k_{cat}/K_m (mM^{-1} s^{-1})	2286.4	923	1580.0	569.16
Guaiacol	K_m (mM)	2.6 ± 0.4	1.16 ± 0.03	6.7 ± 0.7	8.67 ± 0.25
	k_{cat} (s^{-1})	117.4 ± 6.2	65.9 ± 0.5	140.0 ± 5.0	37.15 ± 0.47
	k_{cat}/K_m (mM^{-1} s^{-1})	44.6	57	21.0	4.28

***** Mate et al. 2013

In sum, by ancestral resurrection we have obtained a fungal laccase with higher secretion levels and similar catalytic constants as the evolved OB-1 variant plus added properties in terms of pH stability, yet without passing through the time-consuming process of iterative random mutation, recombination and screening.

4.2.2 Directed evolution of Ancestral Laccase

As a proof of concept that LacAnc100 could be moved towards new directions, we set out the basis for structure-guided evolution towards the oxidation of β-diketones, an engineering goal never tried before. β-diketones are a special class of redox mediators known as initiators which upon oxidation by the laccase they spark the polymerization of vinyl monomers, being part of the final polymeric structures of polystyrene or polyacrylamide (Hollmann and Arends 2012; Hollmann et al. 2008; Durand et al. 2000; Singh et al. 2000; Teixeira et al. 1999). Although using β-diketones to initiate vinyl polymerization is an attractive alternative to replace aggressive chemical methods, laccases hardly oxidize them (Hollmann et al. 2008; Lizotte and Long 2003; Durand et al. 2001). We carefully examined the β-diketone family of compounds, choosing the 1,3 cyclopentanedione for designing a HTS assay for laboratory evolution. This yellow, soluble molecule shows a maximum of absorbance at 460 nm, which upon oxidation by laccase turns to orange, with the concomitant change to 450 nm. The HTS assay was validated by establishing the linearity and the coefficient of variance (15%) of the method from fresh

supernatants produced in microcultures, **Figure 4.2.7**. Two consecutive rescreenings were performed to rule out false positives.

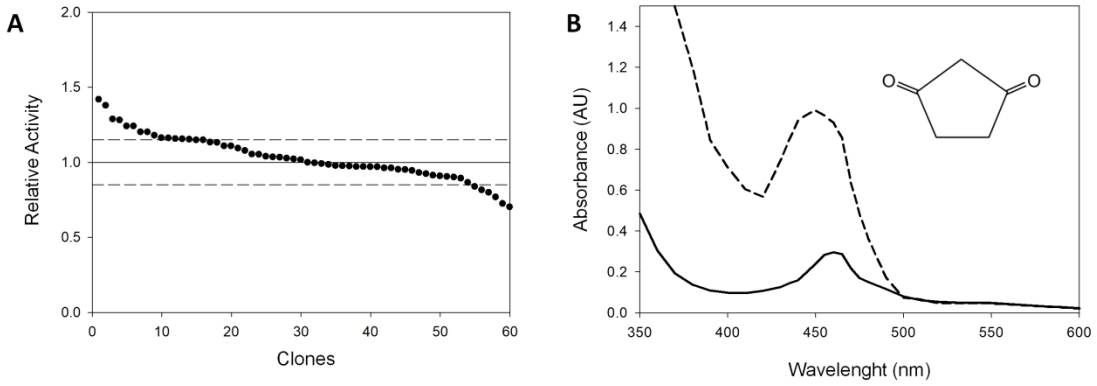


Figure 4.2.7 Screening assay design based on 1,3-cyclopentanedione. **(A)** Activities of the ancestral enzyme plotted in descending order for the 1,3-cyclopentanedione assay. *S. cerevisiae* cells were transformed with PjRoC30-LacAnc100 and plated on SC plates. Individual colonies were picked and inoculated in a 96 well-plate. The activities of the clones were evaluated from fresh supernatant preparations. Dashed lines indicate the coefficient of variation of the assay. **(B)** UV-VIS spectrum of 1,3-cyclopentanedione in the presence of laccase supernatants. Reaction mixture at t=0 (solid line) and t=40h (dashed line). Maximum absorbance peak at 450 nm

We carried out structure-guided evolution by mapping potential residues involved in binding the β-diketone at the T1Cu site of the laccase. Pro163, Val165 and Ile265 were chosen for site saturation mutagenesis. We prepared two independent mutant libraries, saturating by couples Pro163-Ile265 and Val165-Ile265, respectively. The best variants from both libraries were further engineered by site-directed recombination *in vivo* (Viña-Gonzalez et al. 2019) giving rise to the variant P163R-V165R, **Figure 4.2.8 and Figure 3.4.3 (from Materials and Methods section)**. This double mutant showed oxidation rates for 1,3 cyclopetanedione of 160% over parent LacAnc100, being a promising departure point for the future engineering of a competent laccase towards β-diketones.

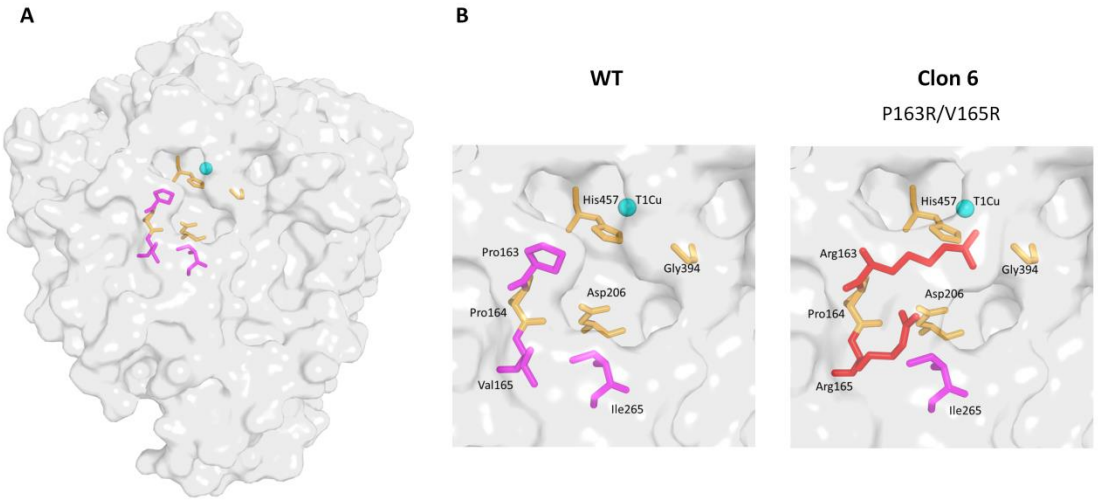


Figure 4.2.8 Molecular model of LacAnc100 laccase including mutations from saturation mutagenesis experiments. (A) The LacAnc100 structure is shown as a light grey surface with some of the residues implicated in substrate interaction represented as sticks. **(B)** Some of the residues of the catalytic pocket are depicted in orange while the selected residues for saturation mutagenesis are highlighted in magenta (left) and the substituted residues in clone 6 are shown in red (right). T1 cooper is represented as a blue sphere. The model was made using Phyre2 server (Protein Homology/analogY Recognition Engine V 2.0: Kelley and Sternberg, 2009) available at www.sbg.bio.ic.ac.uk/phyre2. The outcome model was used for modeling the mutants by Pymol (Schrodinger, LLC [http://www.pymol.org])

4.3 Chapter III: Consensus design of an evolved high-redox potential laccase.

Among the broad repertory of protein engineering methods aimed at improving stability, consensus design has proved to be a powerful strategy for enzyme stabilization without compromising catalytic activity. Here, we have applied a homemade consensus method to stabilize a laboratory evolved high-redox potential laccase. Multiple sequence alignments were carried out and computationally refined by applying relative entropy and mutual information thresholds. With this approach, an ensemble of 20 consensus mutations were identified, 18 of which were ancestral consensus mutations. The set of consensus variants was produced in *S. cerevisiae* and analyzed individually whereas the site directed recombination of best mutations did not yield positive epistasis. The best single variant, carrying the consensus-ancestor mutation A240G in the neighborhood of the T2/T3 copper cluster, improved dramatically thermostability, kinetic values and secretion.

This chapter is based on the article: Gomez-Fernandez B.J., Risso V.A., Sanchez-Ruiz J.M. and Alcalde M. (2019b) Consensus design of an evolved high-redox potential laccase. *Frontiers in Bioengineering and Biotechnology*, Submitted.

4.3.1 Consensus mutations design

The departure point of this study is the OB-1 mutant. Using PM1L as query sequence, we first generated three different MSAs that were manually curated and analyzed (see Experimental procedures for details). Deletion of gaps areas with an occupancy degree of less than 50%, of sequences too short or to large, as well as those that were repeated, reduced the original MSAs to 102, 718 and 147 sequences for MSA1, MSA2 and MSA3, respectively. Then, they were further subjected to computational refinement by applying relative entropy (RE) and mutual information (MI) thresholds, (Sullivan et al. 2012; Magliery and Regan 2005), **Figures 3.5.1 and 3.5.2 (from Materials and Methods section)**. RE provides information about the frequency of an amino acid in a given position whereas MI estimates the RE between pairs (Durani and Magliery 2013). The consensus sequences for the three MSAs were calculated by using RE metric, and these sequences were compared to that of the OB-1 laccase variant so that the positions in which the amino acids differed were targeted as *`potential´* consensus mutations. Substitutions with RE values above the average were preselected and from those, we chose only entries comprising within the three MSAs, which gave rise to 31 consensus mutations over the cut-off, **Figure 4.3.1** To consider the covariance, we also applied the MI metric obtaining

valuable information about possible correlations and anticorrelations that may exist between each pair of positions. Filtering with these constraints allowed the 31 mutations to be cut down to 20 mutations so that only those substitutions with RE values over the average (excluding entries with RE > 2.4 (*i.e.* nearly invariant)) and with MI ≤ 0.6 were selected (Sullivan et al. 2012; Durani and Magliery 2013), **Table 4.3.1**.

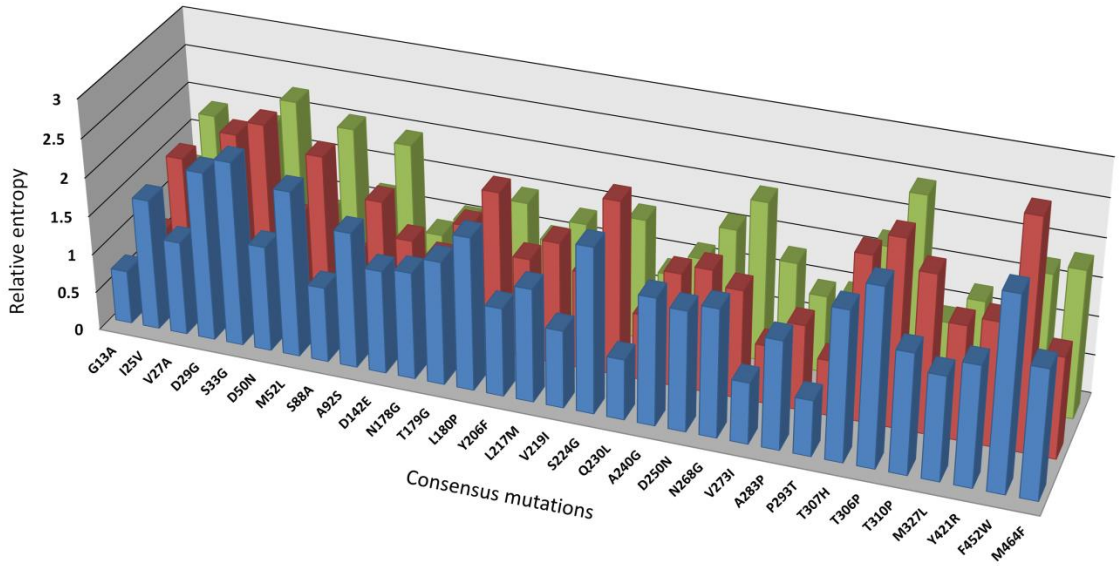


Figure 4.3.1 Relative entropy (RE) distributions for consensus mutations common to the three MSAs. Blue bars, MSA1; Red bars, MSA2; Green bars, MSA3. Despite the different treatment and origin of the sequences from the MSAs, the selected mutations were, in most of the cases, consistent in the RE values for the three groups of alignments.

Consensus mutations often overlap with ancestral mutations -derived from ancestral sequence reconstruction- (Risso et al. 2014). To identify ancestral consensus mutations, we compared the set of selected mutations with the inferred ancestral node of all basidiomycete laccases (LacAnc95) reconstructed for the second chapter of this Doctoral Thesis, dated back ~500 million years (Eastwood 2014; Hedges and Kumar 2009). With the exception of original Val27 and Asn268 which overlapped both in OB-1 and in the ancestral node, the remaining 18 substitutions were ancestral/consensus mutations, **Table 4.3.1**. This data agrees well with previous results in which ancestral sequences are close -but not identical- to consensus counterparts (Risso et al. 2014; Goldsmith and Tawfik 2013 and references herein).

Table 4.3.1 List of consensus mutations that fulfilled the RE and MI thresholds and the comparison with the ancestral node.

Mutation	RE	MI	Consensus/Ancestor
I25V	1.80	0.54	V/V
V27A	1.43	0.47	V/A
S33G	2.37	0.22	G/G
D50N	1.28	0.34	N/N
M52L	2.12	0.14	L/L
A92S	1.70	0.42	S/S
D142E	1.26	0.42	E/E
T179G	1.59	0.53	G/G
Y208F	1.17	0.42	F/F
L217M	1.58	0.42	M/M
V219I	1.26	0.28	I/I
Q230L	0.85	0.51	L/L
A240G	1.47	0.27	G/G
D250N	1.59	0.42	N/N
N268G	1.40	0.58	G/N
T306P	2.40	0.30	P/P
M327L	1.47	0.43	L/L
Y421R	1.59	0.59	R/R
F454W	2.40	0.34	W/W
M464F	1.30	0.29	F/F

4.3.2 Biochemical evaluation of consensus mutations

The 20 consensus mutations were individually inserted into OB-1 laccase, expressed in *S. cerevisiae* and preliminary characterized. Most of the clones were as active as OB-1, while in terms of kinetic thermostability ~1/2 of the variants showed similar or higher values than OB-1, **Figures 4.3.2 and 4.3.3**.

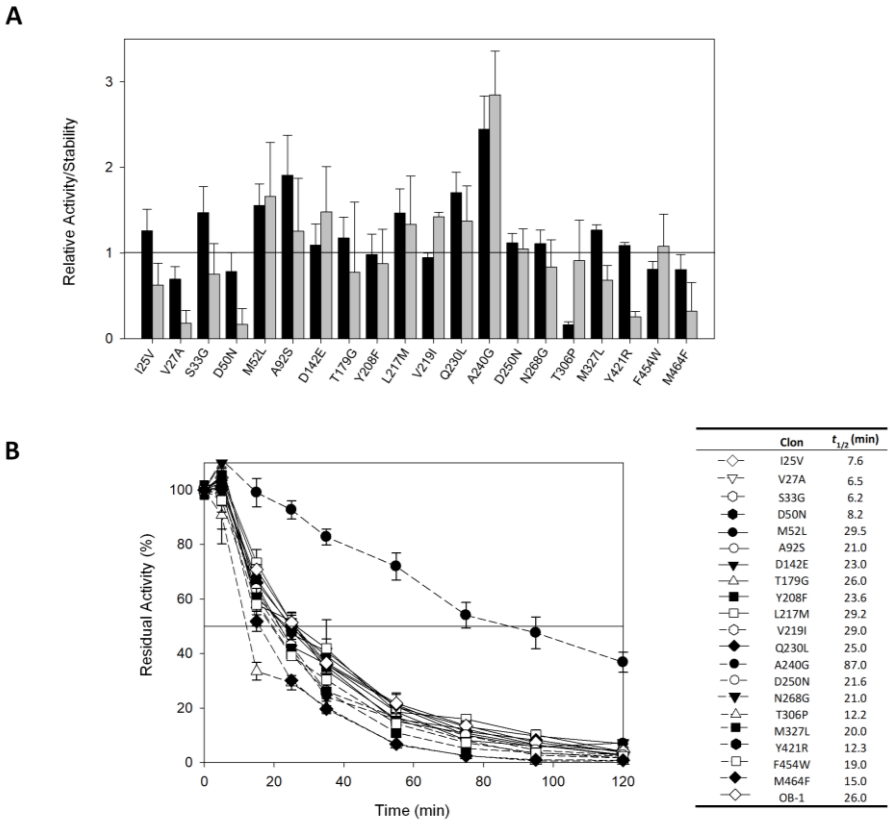


Figure 4.3.2 (A) Evaluation of activities and thermostabilities (in fold-increase) of consensus variants. Solid horizontal line represents the activity and thermostability of the OB1 parental type. Black bars, initial activity; grey bars, stability given as the ratio between residual and initial activities. Laccase activities were normalized to the parental type activity value and each point, including the standard deviation, comes from three independent experiments. **(B)** $t_{1/2}$ (in min) at 70ºC of the twenty consensus variants and OB-1 parental type. Solid horizontal line represents the residual activity at 50%. The values of half-life as well as the symbols that pair with each clone are indicated in a table on the right side of the chart. Each point, including the standard deviation, comes from three independent experiments.

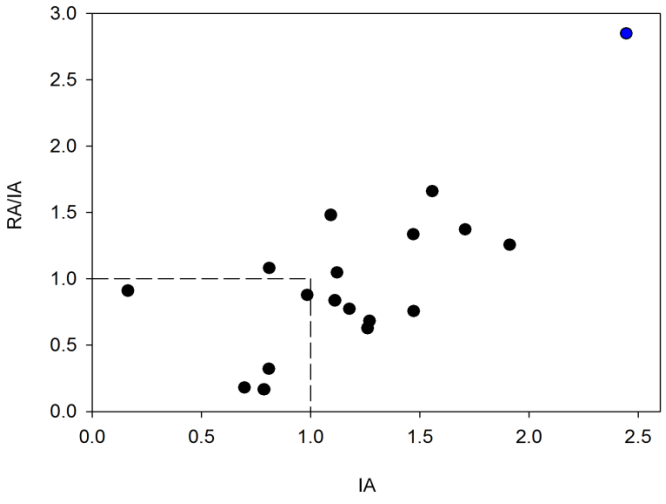


Figure 4.3.3 Functional distribution of consensus variants. IA (initial activity); RA/IA (residual activity/initial activity). Improvements are given in fold vs. OB-1 parental type. The A240G DooKu mutant is highlighted in blue. Dashed horizontal line shows the values of OB-1.

This result is in concordance with previous consensus designs in which roughly 40% to 20% of individual consensus mutations improve thermostability in a greater or lesser extent (Lehmann et al. 2002; Komor et al. 2012; Magliery 2015). As the consensus method tends to break with the stablish trade-off between activity and stability, consensus mutations are usually located far from catalytic sites (Lehmann et al. 2002; Polizzi et al. 2006; Porebski and Buckle 2016; Siddiqui 2017). Accordingly, we mapped the consensus mutations on the model of the OB-1 laccase structure, and with very few exceptions (see below), they were indeed mostly located far from the catalytic copper centers, half of them at the protein surface **Figure 4.3.4; Table 4.3.2**.

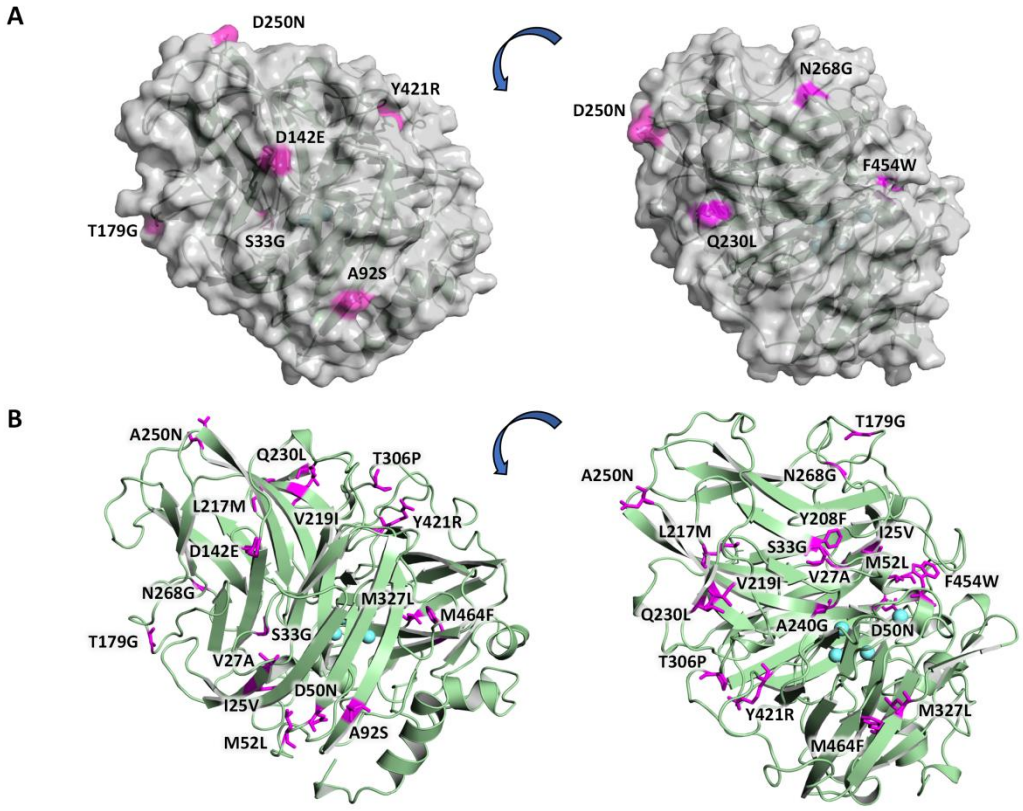


Figure 4.3.4 Location of the consensus mutations in the OB-1 variant. **(A)** Laccase surface is shown in transparent grey while the consensus mutations localized in the surface are highlighted in magenta and labeled. **(B)** Laccase structure shown as green cartoon with the consensus mutations highlighted in magenta and labeled; coopers depicted as blue spheres. Mutations modeled on PDB ID: 5ANH by Pymol (Schrodinger, LLC [http://www.pymol.org]).

Table 4.3.2 Location of consensus mutations.

Mutation	Domain	Structural motif	Relative position	Distance to the T1Cu Site (Å)	Distance to the T2/T3Cu cluster (Å)
I25V	D1	β-sheet	Buried	23.7	15.8
V27A	D1	β-sheet	Buried	25.7	15.6
S33G	D1	Loop	Surface	23.7	14.7
D50N	D1	β-sheet	Partially buried	26.1	18.0
M52L	D1	Loop	Partially buried	24.2	17.6
A92S	D1	β-sheet	Surface	31.8	21.2
D142E	D2	Loop	Surface	29.3	19.7
T179G	D2	Loop	Surface	29.2	29.2
Y208F	D2	β-sheet	Close to Asp205 that interacts with phenols	14.3	13.9
L217M	D2	β-sheet	Partially buried	24.5	22.2
V219I	D2	β-sheet	Buried	23.2	19.1
Q230L	D2	β-sheet	Surface	24.7	23.1
A240G	D2	β-sheet	Close to T2/T3 and His-Cys-His pathway	11.4	5.3
D250N	D2	Loop	Surface	38.0	34.4
N268G	D2	Loop	Surface	23.1	27.0
T306P	D2	Loop	Buried	20.3	25.8
M327L	D3	β-sheet	Partially buried	11.9	13.1
Y421R	D3	Loop	Surface	20.1	13.7
F454W	D3	α-helix	Surface, at the vicinity of the binding pocket	7.3	11.1
M464F	D3	β-sheet	Buried	10.7	17.1

The recombination of consensus mutations may increase thermostability by means of a synergistic process rather than additive (Lehmann et al. 2002; Jones et al. 2017). Taking the most stabilizing consensus mutant as template (A240G -named DooKu-) with an improvement of 61 min in the half-life ($t_{1/2}$) at 70ºC (**Figure 4.3.2B**), we constructed a combinatorial library by site-directed recombination (SDR) *in vivo* (Viña-Gonzalez et al. 2019). With this method, the 6 most promising consensus mutations (M52L, T179G, L217M, V219I, Q230L and D250N) and their corresponding reversions could be rapidly combined on DooKu template, **Figure 3.5.3 (from Materials and Methods section)**. The SDR library was screened for activity and thermostability using a high-throughput assay we developed previously (Garcia-Ruiz et al. 2010), and after two-consecutive rescreenings, 9 variants were selected and preliminary characterized, **Figure 4.3.5 and Figure 3.5.4 (from Materials and Methods section)**.

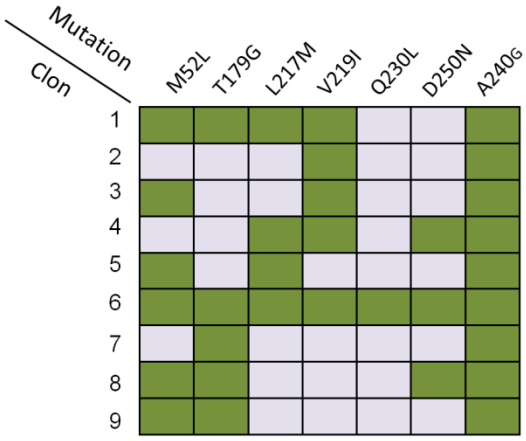
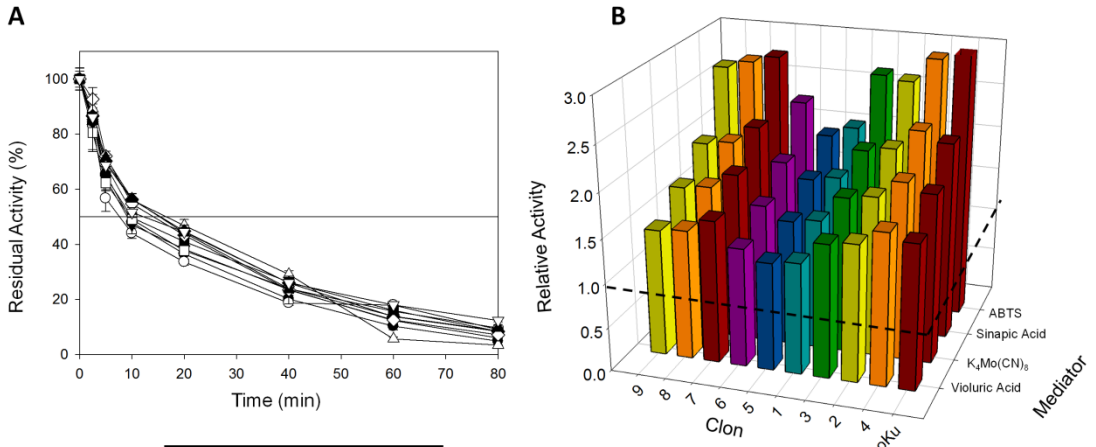


Figure 4.3.5 Site directed recombination *in vivo*. Composition of mutations for the nine recombined variants. Positions mutated to consensus amino acids are highlighted in green squares

Given that the highly stable DooKu mutant was the departure template (with a $t_{1/2}$ at 70ºC of 87 min), for the determination of kinetic thermostability we enhanced the temperature of incubation by 5ºC (*i.e.* $t_{1/2}$ at 75ºC), **Figure 4.3.6A**. The ensemble of recombined variants showed similar thermostabilities as DooKu, in some cases with slight improvements, as in clone 4 which carried mutations L217M, V219I, D250N and A240G, **Figures 4.3.5, 4.3.6A**. Interestingly, the whole set of consensus mutations presented in clone 6 was at the cost of a slight decreased in thermostability, which must be necessarily connected to mutations M52L, T179G and Q230L (absent in clone 4 and exerting negative epistasis) whose RE and MI values alerted us about potential hidden correlations or that may necessitate from other mutations to outweigh. Initial activities were measured with several laccase redox mediators of industrial relevance. In all the cases, the same activity pattern was observed (ABTS > Sinapic acid > $K_4Mo(CN)_8$ > violuric acid) which indicates a strong correlation between the oxidation rate of laccase and the redox potential (*i.e.* the higher the redox potential of the mediator compound the lower the oxidation rate), **Figure 4.3.6B**.



Clon	$t_{1/2}$ (min)
1	12.4
2	10.6
3	12.2
4	16.0
5	9.4
6	9.3
7	12.2
8	11.0
9	12.2
DooKu	11.8

Figure 4.3.6 (A) $t_{1/2}$ (in min) at 75ºC of the nine variants from SDR *in vivo* library and DooKu. Solid dashed line represents the residual activity at 50%. The values of half-life as well as the symbols that pair with each clone are indicated in a table below the chart. Each point, including the standard deviation, comes from three independent experiments. **(B)** Total Activity Improvements (TAI) with four different mediators for the variants from SDR *in vivo* library and DooKu. Dashed horizontal line shows the activity of OB-1. Laccase activities were normalized to the OB-1 activity value and each point, including the standard deviation, comes from three independent experiments.

Given that the most remarkable contribution to activity and stability came from the single mutation A240G, the DooKu variant and the OB-1 parental type were produced at larger scale, purified to homogeneity, and biochemically characterized. Kinetic parameters were evaluated with a selection of redox mediators and phenolic compounds following a similar tendency as that observed with crude supernatants. Irrespective of the substrate, the catalytic efficiency was improved as consequence of an enhanced k_{cat}, **Table 4.3.3**. Indeed, the k_{cat} for DMP, ABTS and sinapic acid was roughly 2-fold over OB-1, while in the case of guaiacol and $K_4Mo(CN)_8$ this increase was smoother. Notably, improved kinetic values were accompanied by a 1.4-fold enhancement in secretion while neither the pH activity profile nor the pH stability varied upon mutation, **Figure 4.3.7**.

Table 4.3.3 Kinetic parameters of OB-1 and DooKu variants. Kinetic constants were estimated in 100 mM phosphate buffer pH 4.0 and 25°C. All reactions were performed by triplicate.

Substrate	Kinetic constant	OB-1	DooKu
ABTS	K_m (mM)	0.007 ± 0.0006	0.009 ± 0.0009
	k_{cat} (s^{-1})	690.0 ± 20.2	1328.8 ± 42.3
	k_{cat}/K_m (mM^{-1} s^{-1})	106073.8	140000.8
DMP	K_m (mM)	0.15 ± 0.01	0.23 ± 0.01
	k_{cat} (s^{-1})	565.0 ± 13	1129.0 ± 18.4
	k_{cat}/K_m (mM^{-1} s^{-1})	3842.6	5561.8
Sinapic Acid	K_m (mM)	0.32 ± 0.02	0.45 ± 0.06
	k_{cat} (s^{-1})	741.0 ± 14.1	1433.06 ± 65.2
	k_{cat}/K_m (mM^{-1} s^{-1})	2286.4	3184.6
$K_4Mo(CN)_8$	K_m (mM)	0.81 ± 0.2	0.81 ± 0.1
	k_{cat} (s^{-1})	220.5 ± 12.1	252.4 ± 11.5
	k_{cat}/K_m (mM^{-1} s^{-1})	271.6	313.2
Guaiacol	K_m (mM)	2.6 ± 0.4	3.4 ± 0.4
	k_{cat} (s^{-1})	117.4 ± 6.2	171.0 ± 7.0
	k_{cat}/K_m (mM^{-1} s^{-1})	44.6	50.3

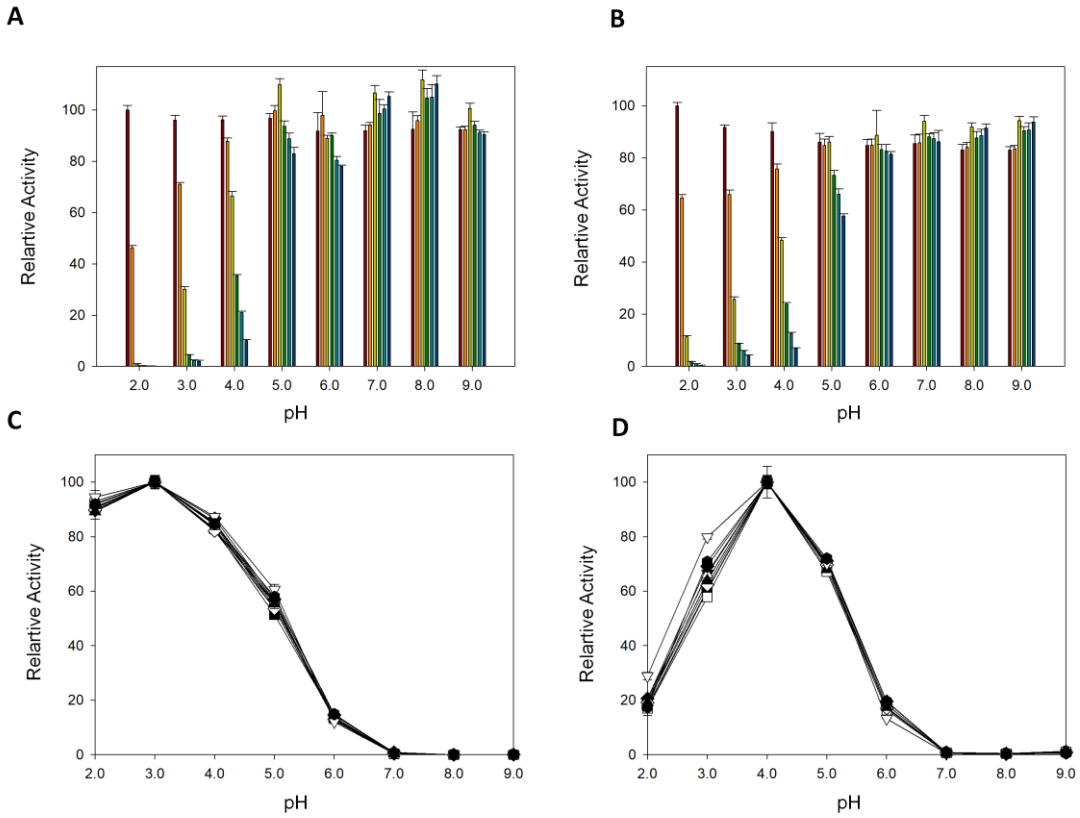


Figure 4.3.7 pH stability of **(A)** the best consensus mutation DooKu and **(B)** the parental type OB-1. Laccases were incubated for 0, 4, 24, 48, 72 and 144 hours at different pHs values (from 2 to 9). Laccase activity was normalized to the highest activity value at time 0. Each point and standard deviation is from three independent measurements. **(C-D)** pH activity profile. Black circles, clon1; White circles, clon2; Black triangles (down), clon 3; White triangles (up), clon4; Black squares, clon5; White squares, clon6; Black diamonds, clon7; White diamonds, clon8; Black triangles (up), clon9; White triangles (down), OB-1; Black circles in dashed lines, DooKu. Activities were measured with ABTS **(C)** or DMP **(D)** as the substrates. Laccase activities were normalized to the optimum activity value and each point, including standard deviation, comes from three different independent experiments.

Consensus mutations that improved both stability and activity are rare since most consensus methods punish positions around the catalytic core (Goldenzweig et al. 2016; Shivange et al. 2016). In the consensus design used in this study, the relaxed RE/MI thresholds allowed us to identify a highly stabilizing consensus/ancestor mutation, located in the surroundings of the catalytic copper cluster T2/T3 for the reduction of O_2 to H_2O. The single change A240G was responsible for a noticeable improvement in thermostability, kinetic parameters and expression albeit the substitution of a hydrophobic Ala by a bulkier Gly did not imply any H-bond interruption or formation, **Figure 4.3.8**. While the striking biochemical effect of this consensus/ancestor mutation is difficult to be rationalized

without entering in speculation, it does prove how consensus design can rapidly unveil hotspot residues for laccase catalysis and function.

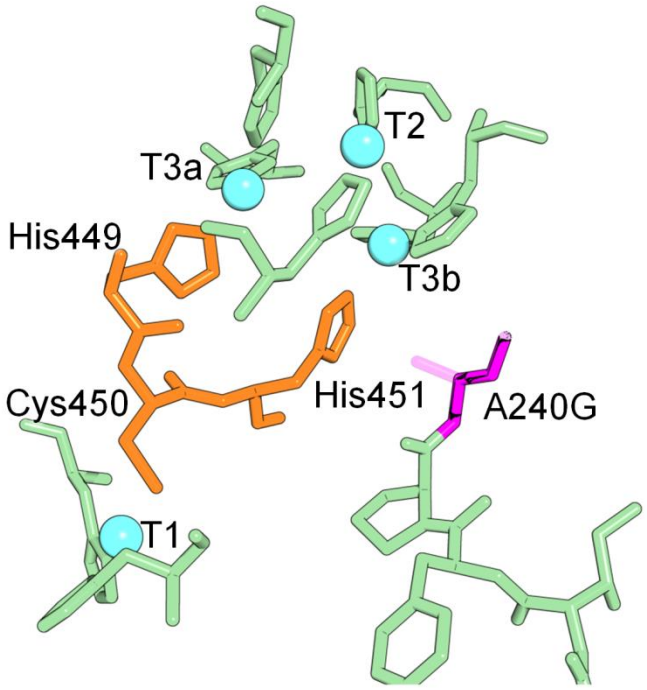


Figure 4.3.8 Detail of the mutation A240G from the DooKu variant in the region of trinuclear Cu cluster. Mutation A240G is shown in magenta with the degraded area corresponding to the lateral chain of the Ala240. The copper atoms are labeled and marked as blue spheres; the residues involved in the transference of electrons from the T1 to the T2/T3 are colored in orange. Residues involved in the first coordination sphere of the catalytic coppers are also represented. Mutations modeled on PDB ID: 5ANH by Pymol (Schrodinger, LLC [http://www.pymol.org]).

Through a homemade consensus method, 20 mutations were predicted and inserted in an evolved HRPL, 18 of which were back-to-ancestor mutations. Roughly half of the mutations incremented thermostability while the rest were neutral and, in a few cases, destabilizing. The best variants were shuffled by site directed recombination, but unlike other consensus designs, no significant positive epistatic effects were found. By contrast, just one single substitution was responsible for a dramatic increase in thermostability, kinetics and secretion.

5 GLOBAL DISCUSSION

5.1 Directed evolution of ancestral enzymes: general considerations

The last decades have witnessed the incremental number of publications related to ancestral sequence reconstruction & resurrection as a rapid methodology to recreate ancestral enzymes (Akanuma 2017; Garcia et al. 2017; Risso et al. 2014; Risso et al. 2013; Perez-Jimenez et al. 2011; Gaucher et al. 2008). Given the appealing characteristics of ancestral enzymes, the potential partnership of ASR and directed evolution has been quickly proposed (Risso et al. 2018; Alcalde 2017b; Merkl and Sterner 2016a; Cox and Gaucher 2014). The first proof of concept combining these two methodologies is now shown within this Doctoral Thesis. Two different modern enzymes, a Rubisco and a laccase from different sources, bacteria and fungi respectively, were selected. The corresponding resurrected enzymes were hypothetically situated far from each other in the evolutionary timeline. While the ancestral Rubiscos were dated in the Precambrian, the ancestral laccases were supposed to appear roughly 2,000 million years later, right after the Cambrian explosion (Hedges et al. 2006). Despite the differences between the two selected nodes (MRPro and LacAnc100), notable similarities spotted during their biochemical characterization and directed evolution.

First, the levels of functional expression for the ancestral enzymes were notably higher than that of extant counterparts. Improved expression of ancestral enzymes is highly convenient, as this is one of the main bottlenecks in enzyme industrial application (Naveen et al. 2019; Rodgers et al. 2010; Martinez et al. 2009). Second, a higher tolerance to mutational loadings was observed for both resurrected enzymes, which correlates with their respective evolution trajectories, with a more relaxed evolution landscape for MRPro than for LacAnc100 -data not shown-. Both properties, strong expression and mutational tolerance, are among the characteristic one might expect from resurrected enzymes (Risso et al. 2018; Alcalde 2017b).

By contrast and unlike other resurrected enzymes, MRPro and LacAnc100 did not show higher thermostability than their modern partners. We only can speculate that MRPro may reside at temperature niches during the cooling of the Precambrian earth (in good agreement with similar results for other Precambrian proteins (Hart et al. 2014, Hobbs et al. 2012) whereas LacAnc100 was `only´ dated back 250 million years, where the

temperature of the planet was similar to that of current earth (Alcalde et al. 2017b and references herein; Hamuro et al. 2017; Eastwood 2014; Gaucher et al. 2008).

In sum, we can conclude that resurrected enzymes are suitable blueprints for laboratory evolution -yet only plausible approximations of true ancestral enzymes-, and that the choice of the enzyme modern template along with the length of the `travel in timeline´ -with different environments among the eras- are important factors to have in mind when starting a joint ASR and directed evolution campaign.

5.1.1 Ancestral Rubisco highlights

Even though Rubisco bridges inorganic and organic phases in the biosphere carbon cycle, becoming responsible for ~95% of the carbon fixed, inherent hurdles to Rubisco catalysis are an extremely poor carboxylation activity and the lack of specificity to discriminate easily between CO_2 and O_2. Together, they make photosynthesis an energetically inefficient process and as such, the Rubisco engineering poses a challenging milestone to increase agriculture productivity and carbon sequestration. The natural trade-off between carboxylation activity and specificity of Rubisco is truly hampering the selection of Rubisco mutants with both improved CO_2 fixation rates and enhanced affinity for CO_2 over O_2, something that strongly connects the advent of Rubisco in our planet to the scarce levels of O_2 in early Precambrian atmosphere, according to predicted geological timescales (Shih et al. 2016; Whitney et al. 2011; Sessions et al. 2009). This supports the natural constraint hypothesized in Rubisco evolution in terms of buffering activity and CO_2 specificity, which soundly explains why all the attempts to evolve in the laboratory superior Rubiscos have led to minor advances. Yet, there are reasons to be optimistic if we observe the higher catalytic behavior of some natural red algae Rubisco (Andrews and Whitney 2003; Whitney et al. 2001), the catalytic discrepancies among different Rubisco forms -even when they all have a highly conserved active site-, or the most recent directed evolution findings assisted by MM1-prk RDE system, from the enhancements in photosynthesis rates >50% in L_8S_8 cyanobacterial Rubisco to the uncoupling of carboxylation activity and CO_2 specificity in L_{10} archaeal Rubisco (Wilson et al. 2016; Durão et al. 2015). Although encouraging, these emergent studies have taught us that the iterative accumulation of just a few beneficial but destabilizing mutations can lead ultimately to the construction of non-functional Rubisco libraries whereby the exploration of the Rubisco sequence space is constrained. Indeed, works on thermophilic archaea Rubisco to enhance activity at ambient temperature (Fujihashi et al. 2016) along with the study of modern and ancestral linages of

thermotolerant cyanobacterial Rubisco from hot springs (Miller et al. 2013) highlight the fundamental role played by stability on Rubisco folding and natural function. In the process of unveiling epistatic effects between mutations, protein evolvability and mutational tolerance perform like communicating vessels so that the capacity to withstand new beneficial mutations during evolution relies on how stable/evolvable the protein scaffold behaves. When the threshold between activity and stability is surpassed, only the insertion of new stabilizing mutations can lead to alternative crossroads in the fitness landscape. RDE system are not suitable to evolve thermostability as they are dependent on low temperatures to slow cell growth, such that Rubisco´s stability can only be sculptured outside of the host biological context (Wilson and Whitney, 2017).

Through the use of an *in vitro* HTS assay, where the trait to be evolved was uncoupled from the biological function, stable Rubisco variants were sorted rapidly. With this method, over 20,000 clones were screened in several campaigns of adaptive, focused evolution and neutral genetic drift subjecting both modern and Precambrian Rubiscos mutant libraries to high temperatures aimed at disclosing a panel of thermostable variants. While Precambrian Rubiscos exhibit most of the properties that one might expect from an ancestral environment (in terms of high solubility and low $S_{C/O}$), they show an unusual mutational tolerance, even in highly conserved catalytic regions, what makes them virgin molds suitable for adaptation. Furthermore, after submitted MRPro to directed evolution in the catalytic pocket, clone B2 spotted with a notable increase in thermostability albeit the cost of a decrease in k_{cat}^{C} that barely affected to the $S_{C/O}$ suggesting the catalytic chemistry of this resurrected mutant diverges significantly from modern Rubiscos where a canonical trade-off between the k_{cat}^{C} and $S_{C/O}$ is observed (Sharwood 2017; Bar-Even et al. 2011; Tcherkez et al. 2006). This is the first proof of concept on predicted Pre-Cambrian enzymes evolved in the laboratory which opens an invaluable pathway to expose the promiscuity and evolvability of ancestral proteins by means of modern directed evolution techniques.

The estimated global population on earth for year 2050 is 9 billion people and 5 more billion people are expected for 2100 addressing as an urgent and paramount challenge the increase of crop productivity (Furbank et al. 2015; Long et al. 2015). Current trends include the design of synthetic photorespiration bypasses and new artificial photorespiration routes in C3 plants, the construction of novel synthetic CO_2-fixation pathways and the engineering of more efficient Rubiscos and CO_2-concentrating mechanisms (Erb and Zarzycki, 2016). The directed *in vitro* evolution platform of modern

and ancestral Rubiscos described in this work complements the abovementioned methods, providing a reliable strategy for selecting highly evolvable, functional Rubiscos.

5.1.2 Ancestral laccase highlights

The broad variety of biotechnological uses of fungal laccases is beyond doubt, as well as the fact that protein engineering (in particular directed evolution) has become the key driver in their adaptation to harsh industrial conditions. Usually, the first premise to tailor laccase, heterologous expression, is an important hurdle that led to a time-consuming process. Indeed, efforts to search universal signal peptides for fungal laccases, have already started along with the set out of several directed evolution platforms for expression (Mateljak et al. 2017; Camarero et al. 2012; Mate et al. 2010, Bulter et al. 2003).

In this Doctoral Thesis we have shown that ASR can bypass laccase expression problems by the resurrection of relatively close ancestors. Of the three laccase nodes, LacAnc100 highlighted in terms of catalytic constants, stability at acid pH and expression.

It is worth noting that it took years to design the OB-1 variant, the product of 8 rounds of directed evolution for expression, stability and activity in yeast. Instead, LacAnc100 was reconstructed and resurrected in a few months, and it showed similar biochemical properties as OB-1.

Finally, the directed evolution of LacAnc100 towards the oxidation of β-diketones is likely to be geared towards a different evolutionary route than OB-1, given that both the ancestral and the modern laccases differ in 134 residues (27% of their protein sequence).

5.2 Ancestral-consensus mutations

ASR and consensus design share the same departure point: to construct a MSA of homologous proteins, yet the rest of the process is different, **Figure 5.2.1**. Interestingly, there are some residues unveiled by ASR that overlap with those from the consensus approach. After all, many consensus mutations belong to ancestral residues strongly conserved, whereas there are some ancestral mutations in regions poorly conserved – *i.e.* that are not consensus-, but precisely the sum of all these residues is the key of the characteristics of ancestral enzymes (Risso et al. 2014; Wijma et al. 2013; Bershtein and Tawfik 2008). In the first two chapters of this Doctoral Thesis we inferred ancestral sequences in the search of new scaffolds to be included within the directed evolution campaigns. In an alternative approach we decided to explore the isolated effect of

consensus-ancestral residues in laccases already evolved in the laboratory, taken out from the equation those non-conserved ancestral residues that could alter the evolved properties. Enzyme engineers obviously wish to improve stability without jeopardizing activity, and the consensus approach has the advantage of overcoming the natural trade-off of activity against stability by inserting stabilizing consensus mutations that do not necessarily dampen activity (Lehmann et al. 2002, Polizzi et al. 2006, Porebski and Buckle 2016).

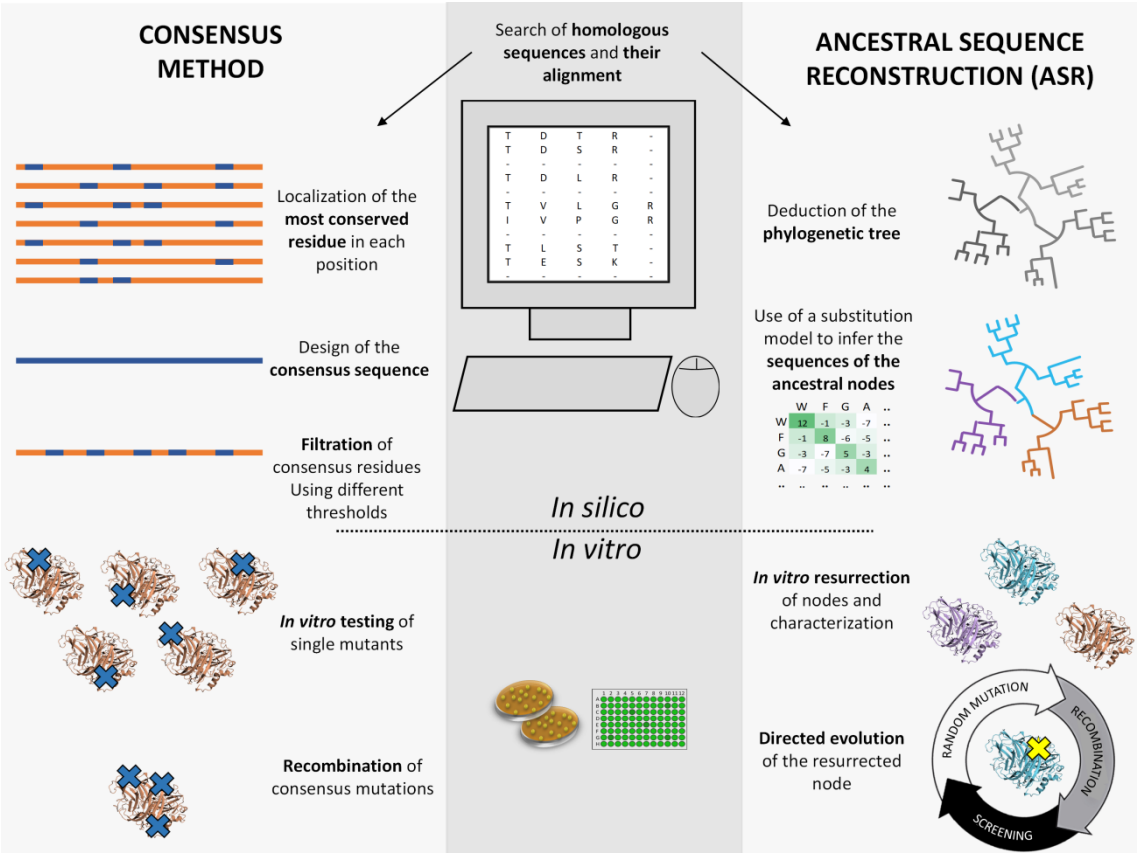


Figure 5.2.1 Comparison between consensus design and ASR. Both methods start from MSAs. **Left:** Consensus sequence is retrieved upon localizing the most conserved residue in each position. Through the application of different metrics, consensus mutations are filtered, *in vitro* evaluated -individually- and/or recombined in the search of positive epistasis. **Right:** Deduction of the phylogenetic tree by using the MSA and Bayesian/ML method. The combination of the sequences, the phylogenetic tree and the substitution model (amino acids or nucleotides) give rise to the reconstructed nodes, which are resurrected (*i.e.* functionally expressed in a modern host), evaluated and further engineered by directed evolution

Due to the nature of consensus mutations, those that strongly stabilized a protein can be beneficial in tailoring homologous enzymes. As such, the best consensus-ancestor mutation A240G could be inserted in other basidiomycete laccases to improve stability without dampening activity. Along these lines, laccase chimeragenesis by SCHEMA-RASPP structure guided recombination *in vivo* could be combined with consensus

mutagenesis and/or ancestral resurrection towards a superior, highly active and stable, laccase variant (Mateljak et al. 2019b).

6 CONCLUSIONS

6.1 Conclusions

Chapter 1

1. Three nodes of bacterial Rubiscos from the Precambrian were obtained through ASR. The most recent common ancestor of all the proteobacteria (2400 million years-myr) MRPro and the most recent common ancestor of β and γ proteobacteria (1900 myr) MRβ/γPro were resurrected (*i.e.* functionally expressed in *E. coli*).
2. The improvements of thermostabilities of the modern Rubisco from *R. rubrum* and for the ancestral MRPro were tackled parallelly by means of different directed evolution strategies. For that purpose, an *in vitro* screening method not linked to the host survival was developed.
3. In addition to a higher mutational tolerance for the insertion of mutations and to higher expression levels for the ancestral counterparts, the evolved variant B2 broke the canonical trade-off between the k_{cat}^{C} and $S_{C/O}$ which opens new venues in the studies aimed at improving these two traits at the same time.

Chapter 2

1. Three nodes of fungal laccases from the Phanerozoic were obtained through ASR. The most recent common ancestor of all Agaricomycotina (430 myr) LacAnc98 and the most recent common ancestor of all Agaricomycetes (252 myr) LacAnc100 were resurrected (*i.e.* functionally expressed in *S. cerevisiae*).
2. The purification and biochemical characterization of LacAnc98 and LacAnc100 showed unusual properties for modern laccases, including a suitable heterologous functional expression and high stability at acid pH.
3. LacAnc100 was evolved towards the oxidation of the initiator 1,3-cyclopentanedione. A high-throughput screening assay was tune up for the oxidation of the initiator which was employed for the focused evolution by combinatorial saturation mutagenesis and site directed recombination *in vivo*. As a result, clone6 spotted (including mutations P163R-V165R) with an improvement of 160% of its activity for 1,3-cyclopentanedione, standing out as departure point for future endeavors.

Chapter 3

1. Through the use of a method that allows to select between different thresholds and ease the user external control, the consensus sequence of a HRPL was calculated.

20 consensus mutations were selected to be included separately *in vitro* in their respective mutants, in order to test independently their impact on the activity and thermostability.

2. The *in vivo* recombination of the best mutations did not render better variants than that that containing the ancestral-consensus A240G mutation (DooKu mutant). The biochemical characterization of DooKu revealed that A240G was not only responsible for a dramatic increment in thermostability (~1 hour in the $t_{1/2}$ at 70ºC) but also for a considerable improvement in the kinetic constants for various substrates.

6.2 Conclusiones

Capítulo 1

i. Mediante ASR se obtuvieron tres nodos de Rubiscos bacterianas situadas en el Precámbrico. Fueron resucitados (*i.e.* expresados funcionalmente en *E. coli*) el ancestro común de las proteobacterias más reciente (2400 Millones de años-Ma) MRPro y el ancestro común de las β y γ proteobacteria más reciente (1900 Ma) MRβ/γPro.

ii. Mediante el uso de diferentes estrategias de evolución dirigida se abordaron de forma paralela las mejoras de las termoestabilidades de la Rubisco moderna de *R. rubrum* y de la ancestral MRPro. Para ello se desarrolló un método de screening *in vitro* no dependiente de la supervivencia del hospedador.

iii. Además de una mayor tolerancia a la introducción de mutaciones y de unos niveles de expresión mejorados para las versiones ancestrales, la variante evolucionada B2 rompió el conservado equilibrio entre $S_{C/O}$ y k_{cat}^{C}, abriendo nuevas posibilidades en los estudios enfocados a mejorar simultáneamente estas dos propiedades.

Capítulo 2

i. Mediante ASR se obtuvieron tres nodos de lacasas fúngicas correspondientes al Fanerozoico. Los ancestros LacAnc98 y LacAnc100 correspondientes al periodo de la aparición del ancestro común de Agaricomycotina y de Agaricomycetes (430-252 Ma), fueron resucitados (*i.e.* expresados funcionalmente en *S. cerevisiae*).

ii. LacAnc98 y LacAnc100 fueron purificadas y caracterizadas bioquímicamente mostrando propiedades no usuales en lacasas fúngicas modernas, que incluyeron una expresión funcional heteróloga apropiada y una alta estabilidad a pH ácido.

iii. El nodo LacAnc100 se evolucionó para la oxidación del iniciador 1,3-ciclopentanodiona. Se puso a punto un método de *high-throughput screening* para la oxidación del iniciador con el que se llevó a cabo la evolución enfocada mediante mutagénesis saturada combinatoria y recombinación dirigida *in vivo*. El resultado de este proceso fue el clon6 (con las mutaciones P163R-V165R), que mejoró un 160% su actividad frente a la 1,3-ciclopentanodiona, siendo un punto de partida prometedor para futuros desarrollos.

Capítulo 3

i. Utilizando un método que permite la aplicación de diferentes umbrales de selección y facilita el control externo del usuario, se calculó la secuencia consenso de una HRPL. Se seleccionaron 20 mutaciones consenso que fueron incluidas individualmente *in vitro* en sus correspondientes mutantes, para evaluar de manera independiente su influencia en la actividad y termoestabilidad.

ii. La recombinación *in vivo* de las mejores mutaciones no generó variantes superiores a la que contenía la mejor mutación consenso-ancestral A240G (mutante DooKu). La caracterización bioquímica de DooKu, reveló que la mutación A240G no sólo fue responsable de un llamativo incremento de termoestabilidad (con un aumento de 1 hora en la $t_{1/2}$ a 70ºC) sino también de una notable mejora en los parámetros cinéticos para diversos sustratos.

7 REFERENCES

Akanuma, S. (2017) Characterization of Reconstructed Ancestral Proteins Suggests a Change in Temperature of the Ancient Biosphere, *Life* **7**: 33.

Akanuma, S., Nakajima, Y., Yokobori, S.-i., Kimura, M., Nemoto, N., Mase, T., et al. (2013) Experimental evidence for the thermophilicity of ancestral life, *Proceedings of the National Academy of Sciences of the United States of America* **110**: 11067-11072.

Alcalde M. (2007) Laccases: Biological Functions, Molecular Structure and Industrial Applications. In: *Industrial Enzymes*. Polaina J., MacCabe A.P. (eds) Springer, Dordrecht.

Alcalde, M. (2010) Mutagenesis protocols in *Saccharomyces cerevisiae* by *in vivo* overlap extension. *Methods in Molecular Biology* **634**: 3-14.

Alcalde, M. (2015) Engineering the ligninolytic enzyme consortium, *Trends in Biotechnology* **33**: 155-162.

Alcalde, M. (2017a) *Directed Enzyme Evolution: Advances and Applications*: Springer.

Alcalde, M. (2017b) When directed evolution met ancestral enzyme resurrection, *Microbial Biotechnology* **10**: 22-24.

Alcalde, M., Bulter, T., and Arnold, F.H. (2002) Colorimetric Assays for Biodegradation of Polycyclic Aromatic Hydrocarbons by Fungal Laccases, *Journal of Biomolecular Screening* **7**: 547-553.

Alcalde, M., Ferrer, M., Plou, F.J., and Ballesteros, A. (2006) Environmental biocatalysis: from remediation with enzymes to novel green processes, *Trends in Biotechnology* **24**: 281-287.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, *Nucleic acids research* **25**: 3389-3402.

Anbar, M., Gul, O., Lamed, R., Sezerman, U.O., and Bayer, E.A. (2012) Improved thermostability of *Clostridium thermocellum* endoglucanase Cel8A by using consensus-guided mutagenesis, *Applied and environmental microbiology* **78**: 3458-3464.

Andersson, I. (2008) Catalysis and regulation in Rubisco, *Journal of experimental botany* **59**: 1555-1568.

Andersson, I., and Taylor, T.C. (2003) Structural framework for catalysis and regulation in ribulose-1, 5-bisphosphate carboxylase/oxygenase, *Archives of Biochemistry and Biophysics* **414**: 130-140.

Andersson, I., and Backlund, A. (2008) Structure and function of Rubisco, *Plant Physiology and Biochemistry* **46**: 275-291.

Andrews, T.J. (1988) Catalysis by cyanobacterial ribulose-bisphosphate carboxylase large subunits in the complete absence of small subunits, *The Journal of Biological Chemistry* **263**: 12213-12219.

Andrews, T.J., and Whitney, S.M. (2003) Manipulating ribulose bisphosphate carboxylase/oxygenase in the chloroplasts of higher plants, *Archives of Biochemistry and Biophysics* **414**: 159-169.

Antonovsky, N., Gleizer, S., and Milo, R. (2017) Engineering carbon fixation in *E. coli*: from heterologous RuBisCO expression to the Calvin–Benson–Bassham cycle, *Current Opinion in Biotechnology* **47**: 83-91.

Applebaum, D. (1996) *Probability and information: An integrated approach*: Cambridge University Press.

Arnold, F.H. (2018) Directed Evolution: Bringing New Chemistry to Life, *Angewandte Chemie (International ed. in English)* **57**: 4143-4148.

Ayuso-Fernandez, I., Martinez, A.T., and Ruiz-Dueñas, F.J. (2017) Experimental recreation of the evolution of lignin-degrading enzymes from the Jurassic to date, *Biotechnology for biofuels* **10**: 67.

Babkova, P., Sebestova, E., Brezovsky, J., Chaloupkova, R., and Damborsky, J. (2017) Ancestral Haloalkane Dehalogenases Show Robustness and Unique Substrate Specificity, *Chembiochem* **18**: 1448-1456.

Baldrian, P. (2006) Fungal laccases – occurrence and properties, *FEMS Microbiology Reviews* **30**: 215-242.

Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D.S., and Milo, R. (2011) The Moderately Efficient Enzyme: Evolutionary and Physicochemical Trends Shaping Enzyme Parameters, *Biochemistry* **50**: 4402-4410.

Bazinet, A.L., Zwickl, D.J., and Cummings, M.P. (2014) A gateway for phylogenetic analysis powered by grid computing featuring GARLI 2.0, *Systematic biology* **63**: 812-818.

Benson, A.A., and Calvin, M. (1948) The path of carbon in photosynthesis. III. In: *Cold Spring Harbor Symposia on Quantitative Biology*: Cold Spring Harbor Laboratory Press. 6-10.

Bershtein, S., Goldin, K., and Tawfik, D.S. (2008) Intense Neutral Drifts Yield Robust and Evolvable Consensus Proteins, *Journal of Molecular Biology* **379**: 1029-1044.

Bershtein, S., and Tawfik, D.S. (2008) Ohno's Model Revisited: Measuring the Frequency of Potentially Adaptive Mutations under Various Mutational Drifts, *Molecular Biology and Evolution* **25**: 2311-2318.

Blatt, L.M., Davis, J.M., Klein, S.B., and Taylor, M.W. (1996) The Biologic Activity and Molecular Characterization of a Novel Synthetic Interferon-Alpha Species, Consensus Interferon, *Journal of Interferon & Cytokine Research* **16**: 489-499.

Bloom, J.D., Labthavikul, S.T., Otey, C.R., and Arnold, F.H. (2006) Protein stability promotes evolvability, *Proceedings of the National Academy of Sciences of the United States of America* **103**: 5869-5874.

Bloom, J.D., Romero, P.A., Lu, Z., and Arnold, F.H. (2007) Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution, *Biology direct* **28**: 2-17.

Bloom, J.D., and Arnold, F.H. (2009) In the light of directed evolution: pathways of adaptive protein evolution, *Proceedings of the National Academy of Sciences of the United States of America* **106** Suppl **1**: 9995-10000.

Bornscheuer, U., and Buchholz, K. (2005) Highlights in biocatalysis–historical landmarks and current trends, *Engineering in life sciences* **5**: 309-323.

Bornscheuer, U., Huisman, G., Kazlauskas, R., Lutz, S., Moore, J., and Robins, K. (2012) Engineering the third wave of biocatalysis, *Nature* **485**: 185-194.

Brijwani, K., Rigdon, A., and Vadlani, P.V. (2010) Fungal laccases: production, function, and applications in food processing, *Enzyme research* **2010**: 1-10.

Bulter, T., Alcalde, M., Sieber, V., Meinhold, P., Schlachtbauer, C., and Arnold, F.H. (2003) Functional expression of a fungal laccase in *Saccharomyces cerevisiae* by directed evolution, *Applied and Environmental Microbiology* **69**: 987-995.

Cai, Z., Liu, G., Zhang, J., and Li, Y. (2014) Development of an activity-directed selection system enabled significant improvement of the carboxylation efficiency of Rubisco, *Protein & cell* **5**: 552-562.

Call, H.P., and Mücke, I. (1997) History, overview and applications of mediated lignolytic systems, especially laccase-mediator-systems (Lignozym®-process), *Journal of Biotechnology* **53**: 163-202.

Camarero, S., Pardo, I., Cañas, A.I., Molina, P., Record, E., Martinez, A.T., et al. (2012) Engineering platforms for directed evolution of Laccase from *Pycnoporus cinnabarinus*, *Applied and Environmental Microbiology* **78**: 1370-1384.

Cambria, M.T., Gullotto, D., Garavaglia, S., and Cambria, A. (2012) *In silico* study of structural determinants modulating the redox potential of *Rigidoporus lignosus* and other fungal laccases, *Journal of Biomolecular Structure and Dynamics* **30**: 89-101.

Campbell, E., Kaltenbach, M., Correy, G.J., Carr, P.D., Porebski, B.T., Livingstone, E.K., et al. (2016) The role of protein dynamics in the evolution of new enzyme function, *Nature chemical biology* **12**: 944-950.

Cañas, A.I., and Camarero, S. (2010) Laccases and their natural mediators: biotechnological tools for sustainable eco-friendly processes, *Biotechnology Advances* **28**: 694-705.

Candel, A.M., Romero-Romero, M.L., Gamiz-Arco, G., Ibarra-Molero, B., and Sanchez-Ruiz, J.M. (2017) Fast folding and slow unfolding of a resurrected Precambrian protein, *Proceedings of the National Academy of Sciences of the United States of America* **114**: E4122-E4123.

Cannatelli, M.D., and Ragauskas, A.J. (2017) Two Decades of Laccases: Advancing Sustainability in the Chemical Industry, *The Chemical Record* **17**: 122-140.

Carmo-Silva, E., Scales, J.C., Madgwick, P.J., and Parry, M.A. (2015) Optimizing Rubisco and its regulation for greater resource use efficiency, *Plant, Cell & Environment* **38**: 1817-1832.

Case, B.A., and Hackel, B.J. (2016) Synthetic and natural consensus design for engineering charge within an affibody targeting epidermal growth factor receptor, *Biotechnology and bioengineering* **113**: 1628-1638.

Castresana, J. (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis, *Molecular Biology and Evolution* **17**: 540-552.

Claus, H. (2004) Laccases: structure, reactions, distribution, *Micron* **35**: 93-96.

Chauhan, P.S., Goradia, B., and Saxena, A. (2017) Bacterial laccase: recent update on production, properties and industrial applications, *3 Biotech* **7**: 323.

Chen, F., Gaucher, E.A., Leal, N.A., Hutter, D., Havemann, S.A., Govindarajan, S., et al. (2010) Reconstructed evolutionary adaptive paths give polymerases accepting reversible terminators for sequencing and SNP detection, *Proceedings of the National Academy of Sciences of the United States of America* **107**: 1948-1953.

Chica, R.A., Doucet, N., and Pelletier, J.N. (2005) Semi-rational approaches to engineering enzyme activity: combining the benefits of directed evolution and rational design, *Current opinion in biotechnology* **16**: 378-384.

Cleland, W.W., Andrews, T.J., Gutteridge, S., Hartman, F.C., and Lorimer, G.H. (1998) Mechanism of Rubisco: the carbamate as general base, *Chemical reviews* **98**: 549-562.

Cliffe, S., Fawer, M.S., Maier, G., Takata, K., and Ritter, G. (1994) Enzyme assays for the phenolic content of natural juices, *Journal of Agricultural and Food Chemistry* **42**: 1824-1828.

Clifton, B.E., Whitfield, J.H., Sanchez-Romero, I., Herde, M.K., Henneberger, C., Janovjak, H., and Jackson, C.J. (2017) Ancestral Protein Reconstruction and Circular Permutation for Improving the Stability and Dynamic Range of FRET Sensors. In: *Synthetic Protein Switches: Methods and Protocols*. Stein, V. (ed). New York, NY: Springer New York. 71-87.

Cline, J., Braman, J.C., and Hogrefe, H.H. (1996) PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases, *Nucleic acids research* **24**: 3546-3551.

Cline, J., and Hogrefe, H. (2000) Randomize gene sequences with new PCR mutagenesis kit, *Strategies* **13**: 157-161.

Cobb, R.E., Chao, R., and Zhao, H. (2013) Directed Evolution: Past, Present and Future, *AIChE journal. American Institute of Chemical Engineers* **59**: 1432-1440.

Cole, M.F., and Gaucher, E.A. (2011) Utilizing natural diversity to evolve protein function: applications towards thermostability, *Current Opinion in Chemical Biology* **15**: 399-406.

Coll, P.M., Fernandez-Abalos, J.M., Villanueva, J.R., Santamaria, R., and Perez, P. (1993) Purification and characterization of a phenoloxidase (laccase) from the lignin-degrading basidiomycete PM1 (CECT 2971), *Applied and Environmental Microbiology* **59**: 2607-2613.

Coordinators, N.R. (2018) Database resources of the National Center for Biotechnology Information, *Nucleic acids research* **46**: D8-D13.

Copley, S.D. (2015) An evolutionary biochemist's perspective on promiscuity, *Trends in biochemical sciences* **40**: 72-78.

Copley, S.D. (2017) Shining a light on enzyme promiscuity, *Current opinion in structural biology* **47**: 167-175.

Cox, V.E., and Gaucher, E.A. (2014) Engineering Proteins by Reconstructing Evolutionary Adaptive Paths. In: *Directed Evolution Library Creation: Methods and Protocols*. Gillam, E.M.J., Copp, J.N., and Ackerley, D. (eds). New York, NY: Springer New York. 353-363.

Cronbach, L.J. (1955) On the non-rational application of information measures in psychology, *Information Theory in Psychology Problems and Methods; Quastler, H., Ed*: 14-30.

Dai, M., Fisher, H.E., Temirov, J., Kiss, C., Phipps, M.E., Pavlik, P., et al. (2007) The creation of a novel fluorescent protein by guided consensus engineering, *Protein Engineering, Design and Selection* **20**: 69-79.

De Groot, A.S., and Scott, D.W. (2007) Immunogenicity of protein therapeutics, *Trends in Immunology* **28**: 482-490.

De Groot, A.S., Ardito, M., Moise, L., Gustafson, E.A., Spero, D., Tejada, G., and Martin, W. (2011) Immunogenic Consensus Sequence T helper Epitopes for a Pan-Burkholderia Biodefense Vaccine, *Immunome research* **7**: e7.

Derango, R.A., Chiang, L.-c., Dowbenko, R., and Lasch, J.G. (1992) Enzyme-mediated polymerization of acrylic monomers, *Biotechnology techniques* **6**: 523-526.

Di Fusco, M., Tortolini, C., Deriu, D., and Mazzei, F. (2010) Laccase-based biosensor for the determination of polyphenol index in wine, *Talanta* **81**: 235-240.

Duff, A.P., Andrews, T.J., and Curmi, P.M. (2000) The transition between the open and closed states of rubisco is triggered by the inter-phosphate distance of the bound bisphosphate, *Journal of molecular biology* **298**: 903-916.

Durand, A., Lalot, T., Brigodiot, M., and Marechal, E. (2000) Enzyme-mediated initiation of acrylamide polymerization: reaction mechanism, *Polymer* **41**: 8183-8192.

Durand, A., Lalot, T., Brigodiot, M., and Marechal, E. (2001) Enzyme-mediated radical initiation of acrylamide polymerization: Main characteristics of molecular weight control, *Polymer* **42**: 5515-5521.

Durani, V. and Magliery, T. J. (2013) Protein engineering and stabilization from sequence statistics: variation and covariation analysis. *Methods in Enzymology* **523**: 237-256.

Durão, P., Bento, I., Fernandes, A.T., Melo, E.P., Lindley, P.F., and Martins, L.O. (2006) Perturbations of the T1 copper site in the CotA laccase from *Bacillus subtilis*: structural, biochemical, enzymatic and stability studies, *JBIC Journal of Biological Inorganic Chemistry* **11**: 514-526.

Durão, P., Aigner, H., Nagy, P., Mueller-Cajar, O., Hartl, F.U., and Hayer-Hartl, M. (2015) Opposing effects of folding and assembly chaperones on evolvability of Rubisco, *Nature Chemical Biology* **11**: 148-155.

Eastwood, D.C. (2014) Evolution of Fungal Wood Decay. In: *Deterioration and Protection of Sustainable Biomaterials*. Schultz T.P., Goodell B., Nicholas D.D. (eds) American Chemical Society. 93-112.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic acids research* **32**: 1792-1797.

Eisenhut, M., Ruth, W., Haimovich, M., Bauwe, H., Kaplan, A., and Hagemann, M. (2008) The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants, *Proceedings of the National Academy of Sciences of the United States of America* **105**: 17199-17204.

Ellis, R.J. (1979) The most abundant protein in the world, *Trends in biochemical sciences* **4**: 241-244.

Endelman, J.B., Silberg, J.J., Wang, Z.-G., and Arnold, F.H. (2004) Site-directed protein recombination as a shortest-path problem, *Protein Engineering, Design and Selection* **17**: 589-594.

Erb, T.J., and Zarzycki, J. (2016) Biochemical and synthetic biology approaches to improve photosynthetic CO_2-fixation, *Current opinion in chemical biology* **34**: 72-79.

Evans, J.R. (2013) Improving Photosynthesis, *Plant Physiology* **162**: 1780-1793.

Feng, X., Berka, R.M., Wahleithner, J.A., Nelson, B.A., Shuster, J.R., Brown, S.H., et al. (1998) Site-directed mutations in fungal laccase: effect on redox potential, activity and pH profile, *Biochemical Journal* **334**: 63-70.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., et al. (2016) The Pfam protein families database: towards a more sustainable future, *Nucleic acids research* **44**: D279-D285.

Fujihashi, M., Nishitani, Y., Kiriyama, T., Aono, R., Sato, T., Takai, T., et al. (2016) Mutation design of a thermophilic Rubisco based on three-dimensional structure enhances its activity at ambient temperature, *Proteins: Structure, Function, and Bioinformatics* **84**: 1339-1346.

Galli, C., and Gentili, P. (2004) Chemical messengers: mediated oxidations with the enzyme laccase, *Journal of Physical Organic Chemistry* **17**: 973-977.

Galli, C., Gentili, P., Jolivalt, C., Madzak, C., and Vadalà, R. (2011) How is the reactivity of laccase affected by single-point mutations? Engineering laccase for improved activity towards sterically demanding substrates, *Applied Microbiology and Biotechnology* **91**: 123-131.

Garcia, A.K., Schopf, J.W., Yokobori, S.-i., Akanuma, S., and Yamagishi, A. (2017) Reconstructed ancestral enzymes suggest long-term cooling of Earth's photic zone since the Archean, *Proceedings of the National Academy of Sciences* **114**: 4619-4624.

Garcia, A.K., and Kaçar, B. (2019) How to resurrect ancestral proteins as proxies for ancient biogeochemistry, *Free Radical Biology and Medicine*.

Garcia-Ruiz, E., Mate, D., Ballesteros, A., Martinez, A. T. and Alcalde, M. (2010) Evolving thermostability in mutant libraries of ligninolytic oxidoreductases expressed in yeast. *Microbial Cell Factories* **9**: 17.

Gaucher, E.A., Gu, X., Miyamoto, M.M., and Benner, S.A. (2002) Predicting functional divergence in protein evolution by site-specific rate shifts, *Trends in Biochemical Sciences* **27**: 315-321.

Gaucher, E.A., Govindarajan, S., and Ganesh, O.K. (2008) Palaeotemperature trend for Precambrian life inferred from resurrected proteins, *Nature* **451**: 704-707.

Georgescu R., Bandara G., Sun L. (2003) Saturation Mutagenesis. In: Arnold F.H., Georgiou G. (eds) *Directed Evolution Library Creation*. Methods in Molecular Biology, vol **231**. Humana Press

Glembo, T.J., Farrell, D.W., Gerek, Z.N., Thorpe, M., and Ozkan, S.B. (2012) Collective dynamics differentiates functional divergence in protein evolution, *PLoS computational biology* **8**: e1002428.

Goldenzweig, A., Goldsmith, M., Hill, S.E., Gertman, O., Laurino, P., Ashani, Y., et al. (2016) Automated structure-and sequence-based design of proteins for high bacterial expression and stability, *Molecular cell* **63**: 337-346.

Goldsmith, M. and Tawfik, D.S. (2013) Enzyme engineering by targeted libraries. *Methods Enzymol* **523**: 257-283.

Gonzalez-Perez, D., Molina-Espeja, P., Garcia-Ruiz, E., and Alcalde, M. (2014) Mutagenic organized recombination process by homologous *in vivo* grouping (MORPHING) for directed enzyme evolution, *PLoS One* **9**: e90919.

Greene, D.N., Whitney, S.M., and Matsumura, I. (2007) Artificially evolved *Synechococcus* PCC6301 Rubisco variants exhibit improvements in folding and catalytic efficiency, *The Biochemical journal* **404**: 517-524.

Guindon, S., and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood, *Systematic biology* **52**: 696-704.

Guo, J., and Zhou, H.-X. (2016) Protein allostery and conformational dynamics, *Chemical reviews* **116**: 6503-6515.

Gupta, R.D., and Tawfik, D.S. (2008) Directed enzyme evolution via small and effective neutral drift libraries, *Nature methods* **5**: 939-942.

Gutteridge, S. (1991) The relative catalytic specificities of the large subunit core of *Synechococcus* ribulose bisphosphate carboxylase/oxygenase, *The Journal of Biological Chemistry* **266**: 7359-7362.

Hadt, R.G., Sun, N., Marshall, N.M., Hodgson, K.O., Hedman, B., Lu, Y., and Solomon, E.I. (2012) Spectroscopic and DFT Studies of Second-Sphere Variants of the Type 1 Copper Site in Azurin: Covalent and Nonlocal Electrostatic Contributions to Reduction Potentials, *Journal of the American Chemical Society* **134**: 16701-16716.

Hamuro, Y., Tajima, K., Matsumoto-Akanuma, A., Sakamoto, S., Furukawa, R., Yamagishi, A., et al. (2017) Characterization of a thermostable mutant of *Agaricus brasiliensis* laccase created by phylogeny-based design, *Journal of bioscience and bioengineering* **124**: 623-629.

Hart, K.M., Harms, M.J., Schmidt, B.H., Elya, C., Thornton, J.W., and Marqusee, S. (2014) Thermodynamic system drift in protein evolution, *PLoS biology* **12**: e1001994.

Hartman, F.C., and Harpel, M.R. (1994) Structure, function, regulation, and assembly of D-ribulose-1, 5-bisphosphate carboxylase/oxygenase, *Annual review of biochemistry* **63**: 197-232.

Hedges, S.B., and Kumar, S. (2009) *The timetree of life*: OUP Oxford.

Heinzelman, P., Romero, P.A., and Arnold, F.H. (2013) Chapter Sixteen - Efficient Sampling of SCHEMA Chimera Families to Identify Useful Sequence Elements. In: *Methods in Enzymology*. Keating, A.E. (ed): Academic Press. 351-368.

Hilgers, R., Vincken, J.-P., Gruppen, H., and Kabel, M.A. (2018) Laccase/Mediator Systems: Their Reactivity toward Phenolic Lignin Structures, *ACS Sustainable Chemistry & Engineering* **6**: 2037-2046.

Hobbs, J.K., Shepherd, C., Saul, D.J., Demetras, N.J., Haaning, S., Monk, C.R., et al. (2011) On the origin and evolution of thermophily: reconstruction of functional Precambrian enzymes from ancestors of *Bacillus*, *Molecular biology and evolution* **29**: 825-835.

Hollmann, F., Gumulya, Y., Tölle, C., Liese, A., and Thum, O. (2008) Evaluation of the Laccase from *Myceliophthora thermophila* as Industrial Biocatalyst for Polymerization Reactions, *Macromolecules* **41**: 8520-8524.

Hollmann, F., and Arends, I.W.C.E. (2012) Enzyme Initiated Radical Polymerizations, *Polymers* **4**.

Hong, G., Ivnitski, D.M., Johnson, G.R., Atanassov, P., and Pachter, R. (2011) Design Parameters for Tuning the Type 1 Cu Multicopper Oxidase Redox Potential: Insight from

a Combination of First Principles and Empirical Molecular Dynamics Simulations, *Journal of the American Chemical Society* **133**: 4802-4809.

Huelsenbeck, J.P., and Bollback, J.P. (2001) Empirical and hierarchical Bayesian estimation of ancestral states, *Systematic biology* **50**: 351-366.

Hullo, M.F., Moszer, I., Danchin, A., and Martin-Verstraete, I. (2001) CotA of *Bacillus subtilis* is a copper-dependent laccase, *Journal of bacteriology* **183**: 5426-5430.

Ikeda, R., Tanaka, H., Uyama, H., and Kobayashi, S. (1998) Laccase-catalyzed polymerization of acrylamide, *Macromolecular rapid communications* **19**: 423-425.

Jackel, C., Bloom, J.D., Kast, P., Arnold, F.H., and Hilvert, D. (2010) Consensus protein design without phylogenetic bias, *Journal of Molecular Biology* **399**: 541-546.

Jensen, R.A. (1976) Enzyme recruitment in evolution of new function, *Annual Review of Microbiology* **30**: 409-425.

Jermann, T.M., Opitz, J.G., Stackhouse, J., and Benner, S.A. (1995) Reconstructing the evolutionary history of the artiodactyl ribonuclease superfamily, *Nature* **374**: 57-59.

Jones, B.J., Lim, H.Y., Huang, J., and Kazlauskas, R.J. (2017) Comparison of Five Protein Engineering Strategies for Stabilizing an alpha/beta-Hydrolase, *Biochemistry* **56**: 6521-6532.

Jones, D.T., Taylor, W.R., and Thornton, J.M. (1992) The rapid generation of mutation data matrices from protein sequences, *Computer Applications in the Biosciences* **8**: 275-282.

Jordan, D.B., and Chollet, R. (1983) Inhibition of ribulose bisphosphate carboxylase by substrate ribulose 1, 5-bisphosphate, *Journal of Biological Chemistry* **258**: 13752-13758.

Kacar, B., Hanson-Smith, V., Adam, Z.R., and Boekelheide, N. (2017) Constraining the timing of the Great Oxidation Event within the Rubisco phylogenetic tree, *Geobiology* **15**: 628-640.

Kaltenbach, M., and Tokuriki, N. (2014) Generation of Effective Libraries by Neutral Drift. In: *Directed Evolution Library Creation: Methods and Protocols*. Gillam, E.M.J., Copp, J.N., and Ackerley, D. (eds). New York, NY: Springer New York. 69-81.

Kallio, J.P., Gasparetti, C., Andberg, M., Boer, H., Koivula, A., Kruus, K., et al. (2011) Crystal structure of an ascomycete fungal laccase from *Thielavia arenaria* – common structural features of asco-laccases, *The FEBS Journal* **278**: 2283-2295.

Kane, H.J., Viil, J., Entsch, B., Paul, K., Morell, M.K., and Andrews, T.J. (1994) An improved method for measuring the CO_2/O_2 specificity of ribulose bisphosphate carboxylase-oxygenase, *Functional Plant Biology* **21**: 449-461.

Kane, H.J., Wilkin, J.-M., Portis, A.R., and Andrews, T.J. (1998) Potent inhibition of ribulose-bisphosphate carboxylase by an oxidized impurity in ribulose-1, 5-bisphosphate, *Plant Physiology* **117**: 1059-1069

Keightley, P.D., and Eyre-Walker, A. (2000) Deleterious Mutations and the Evolution of Sex, *Science* **290**: 331-333.

Kelley, L.A., and Sternberg, M.J. (2009) Protein structure prediction on the Web: a case study using the Phyre server, *Nature protocols* **4**: 363-371.

Kellogg, E.A., and Juliano, N.D. (1997) The structure and function of RuBisCO and their implications for systematic studies, *American journal of botany* **84**: 413-428.

Khersonsky, O., Kiss, G., Röthlisberger, D., Dym, O., Albeck, S., Houk, K.N., et al. (2012) Bridging the gaps in design methodologies by evolutionary optimization of the stability and proficiency of designed Kemp eliminase KE59, *Proceedings of the National Academy of Sciences of the United States of America* **109**: 10358-10363.

Khersonsky, O., Lipsh, R., Avizemer, Z., Ashani, Y., Goldsmith, M., Leader, H., et al. (2018) Automated Design of Efficient and Functionally Diverse Enzyme Repertoires, *Molecular cell* **72**: 178-186.

Kikuchi, M., and Harayama, S. (2002) DNA Shuffling and Family Shuffling for *In vitro* Gene Evolution. In: *In vitro Mutagenesis Protocols*. Braman, J. (ed). Totowa, NJ: Humana Press. 243-257.

Kille, S., Acevedo-Rocha, C.G., Parra, L.P., Zhang, Z.-G., Opperman, D.J., Reetz, M.T., and Acevedo, J.P. (2012) Reducing codon redundancy and screening effort of combinatorial protein libraries created by saturation mutagenesis, *ACS synthetic biology* **2**: 83-92.

Kim, H., Zou, T., Modi, C., Dörner, K., Grunkemeyer, T.J., Chen, L., et al. (2015) A hinge migration mechanism unlocks the evolution of green-to-red photoconversion in GFP-like proteins, *Structure* **23**: 34-43.

Komor, R.S., Romero, P.A., Xie, C.B., and Arnold, F.H. (2012) Highly thermostable fungal cellobiohydrolase I (Cel7A) engineered using predictive methods, *Protein Engineering, Design and Selection* **25**: 827-833.

Koschorreck, K., Schmid, R.D., and Urlacher, V.B. (2009) Improving the functional expression of a *Bacillus licheniformislaccase* by random and site-directed mutagenesis, *BMC Biotechnology* **9**: 12.

Kramer, K.J., Kanost, M.R., Hopkins, T.L., Jing, H., Zhu, Y.C., Xu, R., et al. (2001) Oxidative conjugation of catechols with proteins in insect skeletal systems, *Tetrahedron* **57**: 385-392.

Kratzer, J.T., Lanaspa, M.A., Murphy, M.N., Cicerchi, C., Graves, C.L., Tipton, P.A., et al. (2014) Evolutionary history and metabolic insights of ancient mammalian uricases, *Proceedings of the National Academy of Sciences* **111**: 3763-3768.

Kuchner, O., and Arnold, F.H. (1997) Directed evolution of enzyme catalysts, *Trends in biotechnology* **15**: 523-530.

Kumar, S.V.S., Phale, P.S., Durani, S., and Wangikar, P.P. (2003) Combined sequence and structure analysis of the fungal laccase family, *Biotechnology and Bioengineering* **83**: 386-394.

Kunamneni, A., Camarero, S., Garcia-Burgos, C., Plou, F.J., Ballesteros, A., and Alcalde, M. (2008) Engineering and Applications of fungal laccases for organic synthesis, *Microbial Cell Factories* **7**: 32.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., et al. (2007) Clustal W and Clustal X version 2.0, *Bioinformatics* **23**: 2947-2948.

Lartillot, N., Lepage, T., and Blanquart, S. (2009) PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating, *Bioinformatics* **25**: 2286-2288.

Lee, T.M. (2014) Computationally-guided thermostabilization of the primary endoglucanase from *Hypocrea jecorina* for cellulosic biofuel production. California Institute of Technology.

Lehmann, M., Kostrewa, D., Wyss, M., Brugger, R., D'Arcy, A., Pasamontes, L., and van Loon, A.P. (2000) From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase, *Protein engineering* **13**: 49-57.

Lehmann, M., and Wyss, M. (2001) Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution, *Current Opinion in Biotechnology* **12**: 371-375.

Lehmann, M., Loch, C., Middendorf, A., Studer, D., Lassen, S.F., Pasamontes, L., et al. (2002) The consensus concept for thermostability engineering of proteins: further proof of concept, *Protein engineering* **15**: 403-411.

Letunic, I., Doerks, T., and Bork, P. (2015) SMART: recent updates, new developments and status in 2015, *Nucleic acids research* **43**: 257-260.

Li, H., Webb, S.P., Ivanic, J., and Jensen, J.H. (2004) Determinants of the relative reduction potentials of type-1 copper sites in proteins, *Journal of the American Chemical Society* **126**: 8010-8019.

Li, Y., Drummond, D.A., Sawayama, A.M., Snow, C.D., Bloom, J.D., and Arnold, F.H. (2007) A diverse family of thermostable cytochrome P450s created by recombination of stabilizing fragments, *Nature Biotechnology* **25**: 1051-1056.

Liu, H., Zhu, L., Bocola, M., Chen, N., Spiess, A.C., and Schwaneberg, U. (2013) Directed laccase evolution for improved ionic liquid resistance, *Green Chemistry* **15**: 1348-1355.

Lizotte, J.R., and Long, T.E. (2003) Stable Free-Radical Polymerization of Styrene in Combination with 2-Vinylnaphthalene Initiation, *Macromolecular Chemistry and Physics* **204**: 570-576.

Loening, A.M., Fenn, T.D., Wu, A.M., and Gambhir, S.S. (2006) Consensus guided mutagenesis of *Renilla* luciferase yields enhanced stability and light output, *Protein Engineering, Design and Selection* **19**: 391-400.

Long, Stephen P., Marshall-Colon, A., and Zhu, X.-G. (2015) Meeting the Global Food Demand of the Future by Engineering Crop Photosynthesis and Yield Potential, *Cell* **161**: 56-66.

Magliery, T. J. (2015) Protein stability: computation, sequence statistics, and new experimental methods. *Current Opinion in Structural Biology* **33**: 161-168.

Magliery, T.J., and Regan, L. (2005) Sequence variation in ligand binding sites in proteins, *BMC bioinformatics* **6**: 240.

Manteca, A., Schönfelder, J., Alonso-Caballero, A., Fertin, M.J., Barruetabeña, N., Faria, B.F., et al. (2017) Mechanochemical evolution of the giant muscle protein titin as inferred from resurrected proteins, *Nature structural & molecular biology* **24**: 652-657.

Mark, H.F., McKetta, J.J., and Othmer, D.F. (1963) *Kirk-Othmer encyclopedia of chemical technology*: Interscience.

Marshall, N.M., Garner, D.K., Wilson, T.D., Gao, Y.-G., Robinson, H., Nilges, M.J., and Lu, Y. (2009) Rationally tuning the reduction potential of a single cupredoxin beyond the natural range, *Nature* **462**: 113-116.

Martinez, A.T., Ruiz-Duenas, F.J., Martinez, M.J., del Rio, J.C., and Gutierrez, A. (2009) Enzymatic delignification of plant cell wall: from nature to mill, *Current opinion in biotechnology* **20**: 348-357.

Martins, L.O., Soares, C.M., Pereira, M.M., Teixeira, M., Costa, T., Jones, G.H., and Henriques, A.O. (2002) Molecular and biochemical characterization of a highly stable bacterial laccase that occurs as a structural component of the *Bacillus subtilis* endospore coat, *The Journal of Biological Chemistry* **277**: 18849-18859.

Mate, D., Garcia-Burgos, C., Garcia-Ruiz, E., Ballesteros, A.O., Camarero, S., and Alcalde, M. (2010) Laboratory evolution of high-redox potential laccases, *Chemistry & Biology* **17**: 1030-1041.

Mate, D.M., Garcia-Ruiz, E., Camarero, S., Shubin, V.V., Falk, M., Shleev, S., et al. (2013) Switching from blue to yellow: altering the spectral properties of a high redox potential laccase by directed evolution, *Biocatalysis and Biotransformation* **31**: 8-21.

Mate, D.M., and Alcalde, M. (2015) Laccase engineering: from rational design to directed evolution, *Biotechnology Advances* **33**: 25-40.

Mate, D.M., and Alcalde, M. (2017) Laccase: a multi-purpose biocatalyst at the forefront of biotechnology, *Microbial Biotechnology* **10**: 1457-1467.

Mate, D.M., Gonzalez-Perez, D., Mateljak, I., Gomez de Santos, P., Vicente, A.I., and Alcalde, M. (2017) Chapter 8 - The Pocket Manual of Directed Evolution: Tips and Tricks. In: *Biotechnology of Microbial Enzymes*. Brahmachari, G. (ed): Academic Press. 185-213.

Mateljak, I., Tron, T., and Alcalde, M. (2017) Evolved α-factor prepro-leaders for directed laccase evolution in Saccharomyces cerevisiae, *Microbial biotechnology* **10**: 1830-1836.

Mateljak, I., Monza, E., Lucas, M.F., Guallar, V., Aleksejeva, O., Ludwig, R., et al. (2019a) Increasing Redox Potential, Redox Mediator Activity, and Stability in a Fungal Laccase by Computer-Guided Mutagenesis and Directed Evolution, *ACS Catalysis* **9**: 4561-4572.

Mateljak, I., Rice, A., Yang, K., Tron, T., and Alcalde, M. (2019b) The Generation of Thermostable Fungal Laccase Chimeras by SCHEMA-RASPP Structure-Guided Recombination *in vivo*, *ACS Synthetic Biology* **8**: 833-843.

Matera, I., Gullotto, A., Tilli, S., Ferraroni, M., Scozzafava, A., and Briganti, F. (2008) Crystal structure of the blue multicopper oxidase from the white-rot fungus *Trametes trogii* complexed with p-toluate, *Inorganica Chimica Acta* **361**: 4129-4137.

Mayer, A.M., and Staples, R.C. (2002) Laccase: new functions for an old enzyme, *Phytochemistry* **60**: 551-565.

McCurry, S.D., Pierce, J., Tolbert, N.E., and Orme-Johnson, W.H. (1981) On the mechanism of effector-mediated activation of ribulose bisphosphate carboxylase/oxygenase, *The Journal of Biological Chemistry* **256**: 6623-6628.

Merkl, R., and Sterner, R. (2016a) Ancestral protein reconstruction: techniques and applications, *Biological chemistry* **397**: 1-21.

Merkl, R., and Sterner, R. (2016b) Reconstruction of ancestral enzymes, *Perspectives in Science* **9**: 17-23.

Mikolasch, A., and Schauer, F. (2009) Fungal laccases as tools for the synthesis of new hybrid molecules and biomaterials, *Applied Microbiology and Biotechnology* **82**: 605-624.

Miller, S.R., McGuirl, M.A., and Carvey, D. (2013) The evolution of RuBisCO stability at the thermal limit of photoautotrophy, *Molecular biology and evolution* **30**: 752-760.

Milligan, C., and Ghindilis, A. (2002) Laccase Based Sandwich Scheme Immunosensor Employing Mediatorless Electrocatalysis, *Electroanalysis* **14**: 415-419.

Miyazaki, J., Nakaya, S., Suzuki, T., Tamakoshi, M., Oshima, T., and Yamagishi, A. (2001) Ancestral Residues Stabilizing 3-Isopropylmalate Dehydrogenase of an Extreme Thermophile: Experimental Evidence Supporting the Thermophilic Common Ancestor Hypothesis, *The Journal of Biochemistry* **129**: 777-782.

Molina-Espeja, P., Vina-Gonzalez, J., Gomez-Fernandez, B.J., Martin-Diaz, J., Garcia-Ruiz, E., and Alcalde, M. (2016) Beyond the outer limits of nature by directed evolution, *Biotechnology advances* **34**: 754-767.

Mollania, N., Khajeh, K., Ranjbar, B., and Hosseinkhani, S. (2011) Enhancement of a bacterial laccase thermostability through directed mutagenesis of a surface loop, *Enzyme and Microbial Technology* **49**: 446-452.

Morozova, O.V., Shumakovich, G.P., Gorbacheva, M.A., Shleev, S.V., and Yaropolov, A.I. (2007a) "Blue" laccases, *Biochemistry (Mosc)* **72**: 1136-1150.

Morozova, O.V., Shumakovich, G.P., Shleev, S.V., and Iaropolov, A.I. (2007b) [Laccase-mediator systems and their applications: a review], *Applied Biochemistry and Microbiology* **43**: 583-597.

Mot, A.C., and Silaghi-Dumitrescu, R. (2012) Laccases: Complex architectures for one-electron oxidations, *Biochemistry (Moscow)* **77**: 1395-1407.

Mueller-Cajar, O., Morell, M., and Whitney, S.M. (2007) Directed Evolution of Rubisco in *Escherichia coli* Reveals a Specificity-Determining Hydrogen Bond in the Form II Enzyme, *Biochemistry* **46**: 14067-14074.

Mueller-Cajar, O., and Whitney, S.M. (2008a) Directing the evolution of Rubisco and Rubisco activase: first impressions of a new tool for photosynthesis research, *Photosynthesis research* **98**: 667-675.

Mueller-Cajar, O., and Whitney, S. M. (2008b) Evolving improved *Synechococcus* Rubisco functional expression in *Escherichia coli*, *Biochemical Journal* **414**: 205-214.

Musil, M., Stourac, J., Bendl, J., Brezovsky, J., Prokop, Z., Zendulka, J., et al. (2017) FireProt: web server for automated design of thermostable proteins, *Nucleic Acids Research* **45**: 393-399.

Narancic, T., Davis, R., Nikodinovic-Runic, J., and O' Connor, K.E. (2015) Recent developments in biocatalysis beyond the laboratory, *Biotechnology Letters* **37**: 943-954.

Naveen, P., Shraddha, S., Shivam, Ria, M., and Umesh, M. (2019) Mode of Action, Properties, Production, and Application of Laccase: A Review, *Recent Patents on Biotechnology* **13**: 19-32.

Ness, J.E., Kim, S., Gottman, A., Pak, R., Krebber, A., Borchert, T.V., et al. (2002) Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently, *Nature biotechnology* **20**: 1251-1255.

Nestl, B.M., Nebel, B.A., and Hauer, B. (2011) Recent progress in industrial biocatalysis, *Current opinion in chemical biology* **15**: 187-193.

Nestl B.M., Hauer B.A. (2014) Engineering of flexible loops in enzymes. *ACS Catalysis*.**4**: 3201-3211.

Nguyen, V., Wilson, C., Hoemberger, M., Stiller, J.B., Agafonov, R.V., Kutter, S., et al. (2017) Evolutionary drivers of thermoadaptation in enzyme catalysis, *Science* **355**: 289-294.

Nikolova, P.V., Henckel, J., Lane, D.P., and Fersht, A.R. (1998) Semirational design of active tumor suppressor p53 DNA binding domain with enhanced stability, *Proceedings of the National Academy of Sciences of the United States of America* **95**: 14675-14680.

Nisbet, E.G., Grassineau, N.V., Howe, C.J., Abell, P.I., Regelous, M., and Nisbet, R.E.R. (2007) The age of Rubisco: the evolution of oxygenic photosynthesis, *Geobiology* **5**: 311-335.

Osipov, E., Polyakov, K., Kittl, R., Shleev, S., Dorovatovsky, P., Tikhonova, T., et al. (2014) Effect of the L499M mutation of the ascomycetous *Botrytis aclada* laccase on redox potential and catalytic properties, *Acta crystallographica. Section D, Biological crystallography* **70**: 2913-2923.

Osma, J.F., Toca-Herrera, J.L., and Rodriguez-Couto, S. (2010) Uses of laccases in the food industry, *Enzyme research* **2010**: 918761.

Paatero, A., Rosti, K., Shkumatov, A.V., Sele, C., Brunello, C., Kysenius, K., et al. (2016) Crystal Structure of an Engineered LRRTM2 Synaptic Adhesion Molecule and a Model for Neurexin Binding, *Biochemistry* **55**: 914-926.

Packer, M.S., and Liu, D.R. (2015) Methods for the directed evolution of proteins, *Nature Reviews Genetics* **16**: 379-394.

Pantoliano, M.W., Whitlow, M., Wood, J.F., Dodd, S.W., Hardman, K.D., Rollence, M.L., and Bryan, P.N. (1989) Large increases in general stability for subtilisin BPN' through incremental changes in the free energy of unfolding, *Biochemistry* **28**: 7205-7213.

Pardo, I., Santiago, G., Gentili, P., Lucas, F., Monza, E., Medrano, F.J., et al. (2016) Re-designing the substrate binding pocket of laccase for enhanced oxidation of sinapic acid, *Catalysis Science & Technology* **6**: 3900-3910.

Pardo, I., Rodriguez-Escribano, D., Aza, P., de Salas, F., Martinez, A.T., and Camarero, S. (2018) A highly stable laccase obtained by swapping the second cupredoxin domain, *Scientific Reports* **8**: 15669.

Parry, M., Andralojc, P., Mitchell, R.A., Madgwick, P., and Keys, A. (2003) Manipulation of Rubisco: the amount, activity, function and regulation, *Journal of experimental botany* **54**: 1321-1333.

Parry, M.A., Andralojc, P.J., Scales, J.C., Salvucci, M.E., Carmo-Silva, A.E., Alonso, H., and Whitney, S.M. (2013) Rubisco activity and regulation as targets for crop improvement, *Journal of Experimental Botany* **64**: 717-730.

Pauling L, Zuckerkandl, E. (1963) Molecular "restoration studies" of extinct forms of life, *Acta Chemica Scandinavica* **17**: 9–16.

Pearson, W.R. (2013) An Introduction to Sequence Similarity ("Homology") Searching, *Current Protocols in Bioinformatics* **42**: 3.1.1-3.1.8.

Perez-Jimenez, R., Ingles-Prieto, A., Zhao, Z.-M., Sanchez-Romero, I., Alegre-Cebollada, J., Kosuri, P., et al. (2011) Single-molecule paleoenzymology probes the chemistry of resurrected enzymes, *Nature structural & molecular biology* **18**: 592-596.

Pierce, J., Tolbert, N., and Barker, R. (1980) Interaction of ribulose bisphosphate carboxylase/oxygenase with transition-state analogs, *Biochemistry* **19**: 934-942

Piontek, K., Antorini, M., and Choinowski, T. (2002) Crystal Structure of a Laccase from the Fungus *Trametes versicolor* at 1.90-Å Resolution Containing a Full Complement of Coppers, *Journal of Biological Chemistry* **277**: 37663-37669.

Polizzi, K.M., Chaparro-Riggers, J.F., Vazquez-Figueroa, E., and Bommarius, A.S. (2006) Structure-guided consensus approach to create a more thermostable penicillin G acylase, *Biotechnology Journal* **1**: 531-536.

Porebski, B.T., and Buckle, A.M. (2016) Consensus protein design, *Protein Engineering, Design and Selection* **29**: 245-251.

Portis, A.R. (2003) Rubisco activase–Rubisco's catalytic chaperone, *Photosynthesis research* **75**: 11-27.

Portis, A.R., and Parry, M.A. (2007) Discoveries in Rubisco (Ribulose 1, 5-bisphosphate carboxylase/oxygenase): a historical perspective, *Photosynthesis Research* **94**: 121-143.

Price, G.D., and Howitt, S.M. (2014) Plant science: Towards turbocharged photosynthesis, *Nature* **513**: 497-498.

Priddy, D. (1994) Recent advances in styrene polymerization. In: *Polymer Synthesis*: Springer. 67-114.

Pupko, T., Pe, I., Shamir, R., and Graur, D. (2000) A fast algorithm for joint reconstruction of ancestral amino acid sequences, *Molecular biology and evolution* **17**: 890-896.

Rasekh, B., Khajeh, K., Ranjbar, B., Mollania, N., Almasinia, B., and Tirandaz, H. (2014) Protein engineering of laccase to enhance its activity and stability in the presence of organic solvents, *Engineering in Life Sciences* **14**: 442-448.

Reetz, M.T. (2017) Directed evolution of enzyme robustness. In: *Directed evolution of selective enzymes*: *Catalysis for organic chemistry and biotechnology*. Reetz, M.T. (eds). Wiley-VCH Verlag GmbH & Co. pp. 205-235.

Reetz, M.T., and Carballeira, J.D. (2007) Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes, *Nature Protocols* **2**: 891-903.

Reetz, M.T., Kahakeaw, D., and Lohmer, R. (2008) Addressing the numbers problem in directed evolution, *ChemBioChem* **9**: 1797-1804.

Risso, V.A., Gavira, J.A., Mejia-Carmona, D.F., Gaucher, E.A., and Sanchez-Ruiz, J.M. (2013) Hyperstability and Substrate Promiscuity in Laboratory Resurrections of Precambrian β-Lactamases, *Journal of the American Chemical Society* **135**: 2899-2902.

Risso, V.A., Gavira, J.A., and Sanchez-Ruiz, J.M. (2014) Thermostable and promiscuous Precambrian proteins, *Environmental Microbiology* **16**: 1485-1489.

Risso, V.A., Martinez-Rodriguez, S., Candel, A.M., Krüger, D.M., Pantoja-Uceda, D., Ortega-Muñoz, M., et al. (2017) De novo active sites for resurrected Precambrian enzymes, *Nature Communications* **8**: 16113.

Risso, V.A., and Sanchez-Ruiz, J.M. (2017) Resurrected Ancestral Proteins as Scaffolds for Protein Engineering. In: *Directed Enzyme Evolution: Advances and Applications*. Alcalde, M. (ed). Cham: Springer International Publishing. 229-255.

Risso, V.A., Sanchez-Ruiz, J.M., and Ozkan, S.B. (2018) Biotechnological and protein-engineering implications of ancestral protein resurrection, *Current Opinion in Structural Biology* **51**: 106-115.

Riva, S. (2006) Laccases: blue enzymes for green chemistry, *Trends in Biotechnology* **24**: 219-226.

Robert, F., and Chaussidon, M. (2006) A palaeotemperature curve for the Precambrian oceans based on silicon isotopes in cherts, *Nature* **443**: 969-972.

Rodgers, C.J., Blanford, C.F., Giddens, S.R., Skamnioti, P., Armstrong, F.A., and Gurr, S.J. (2010) Designer laccases: a vogue for high-potential fungal enzymes?, *Trends in Biotechnology* **28**: 63-72.

Rodriguez, F., Cohen, C., Ober, C.K., and Archer, L. (2014) *Principles of polymer systems*: CRC Press.

Ronquist, F., and Huelsenbeck, J.P. (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models, *Bioinformatics* **19**: 1572-1574.

Rost, B. (1999) Twilight zone of protein sequence alignments, *Protein engineering* **12**: 85-94.

Ryden, L.G., and Hunt, L.T. (1993) Evolution of protein complexity: The blue copper-containing oxidases and related proteins, *Journal of Molecular Evolution* **36**: 41-66.

Saab-Rincon, G., Alwaseem, H., Guzman-Luna, V., Olvera, L., and Fasan, R. (2018) Stabilization of the Reductase Domain in the Catalytically Self-Sufficient Cytochrome P450(BM3) by Consensus-Guided Mutagenesis, *Chembiochem : a European journal of chemical biology* **19**: 622-632.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*: Cold Spring Harbor Laboratory Press.

Santiago, G., de Salas, F., Lucas, M.F., Monza, E., Acebes, S., Martinez, A.T., et al. (2016) Computer-Aided Laccase Engineering: Toward Biological Oxidation of Arylamines, *ACS Catalysis* **6**: 5415-5423.

Satagopan, S., Scott, S.S., Smith, T.G., and Tabita, F.R. (2009) A Rubisco mutant that confers growth under a normally "inhibitory" oxygen concentration, *Biochemistry* **48**: 9076-9083.

Sato, T., Atomi, H., and Imanaka, T. (2007) Archaeal Type III RuBisCOs Function in a Pathway for AMP Metabolism, *Science* **315**: 1003-1006.

Scheiblbrandner, S., Breslmayr, E., Csarman, F., Paukner, R., Führer, J., Herzog, P.L., et al. (2017) Evolving stability and pH-dependent activity of the high redox potential *Botrytis aclada* laccase for enzymatic fuel cells, *Scientific reports* **7**: 13688.

Schneider, G., Lindqvist, Y., and Lundqvist, T. (1990) Crystallographic refinement and structure of ribulose-1, 5-bisphosphate carboxylase from *Rhodospirillum rubrum* at 1.7 Å resolution, *Journal of molecular biology* **211**: 989-1008.

Schreuder, H.A., Knight, S., Curmi, P.M., Andersson, I., Cascio, D., Branden, C.I., and Eisenberg, D. (1993) Formation of the active site of ribulose-1,5-bisphosphate carboxylase/oxygenase by a disorder-order transition from the unactivated to the activated form, *Proceedings of the National Academy of Sciences of the United States of America* **90**: 9968-9972.

Sessions, A.L., Doughty, D.M., Welander, P.V., Summons, R.E., and Newman, D.K. (2009) The continuing puzzle of the great oxidation event, *Current biology* **19**: 567-574.

Shannon, C.E., and Weaver, W. (1949) *The mathematical theory of information*. Urbana, IL: The University of Illinois Press, 1-117.

Sharwood, R.E. (2017) Engineering chloroplasts to improve Rubisco catalysis: prospects for translating improvements into food and fiber crops, *New Phytologist* **213**: 494-510.

Sharwood, R.E., Sonawane, B.V., Ghannoum, O., and Whitney, S.M. (2016a) Improved analysis of C4 and C3 photosynthesis via refined *in vitro* assays of their carbon fixation biochemistry, *Journal of experimental botany* **67**: 3137-3148.

Sharwood, R.E., Ghannoum, O., and Whitney, S.M. (2016b) Prospects for improving CO_2 fixation in C3-crops through understanding C4-Rubisco biogenesis and catalytic diversity, *Current Opinion in Plant Biology* **31**: 135-142.

Sheng, S., Jia, H., Topiol, S., and Farinas, E.T. (2017) Engineering CotA Laccase for Acidic pH Stability Using *Bacillus subtilis* Spore Display, *Journal of Microbiology and Biotechnology* **27**: 507-513.

Shih, P.M., Occhialini, A., Cameron, J.C., Andralojc, P.J., Parry, M.A., and Kerfeld, C.A. (2016) Biochemical characterization of predicted Precambrian RuBisCO, *Nature communications* **7**: 10382.

Shin, S.K., Hyeon, J.E., Joo, Y.-C., Jeong, D.W., You, S.K., and Han, S.O. (2019) Effective melanin degradation by a synergistic laccase-peroxidase enzyme complex for skin whitening and other practical applications, *International Journal of Biological Macromolecules* **129**: 181-186.

Shivange, A.V., Marienhagen, J., Mundhada, H., Schenk, A., and Schwaneberg, U. (2009) Advances in generating functional diversity for directed protein evolution, *Current opinion in chemical biology* **13**: 19-25.

Shivange, A.V., Hoeffken, H.W., Haefner, S. and Schwaneberg, U. (2016) Protein consensus-based surface engineering (ProCoS): a computer-assisted method for directed protein evolution. *BioTechniques* **61**(6): 305-314

Shleev, S., Jarosz-Wilkolazka, A., Khalunina, A., Morozova, O., Yaropolov, A., Ruzgas, T., and Gorton, L. (2005) Direct electron transfer reactions of laccases from different origins on carbon electrodes, *Bioelectrochemistry* **67**: 115-124.

Siddiqui, K.S. (2017) Defying the activity-stability trade-off in enzymes: taking advantage of entropy to enhance activity and thermostability. *Critical Reviews in Biotechnology* **37**: 309-322.

Sigrist, C.J.A., de Castro, E., Cerutti, L., Cuche, B.A., Hulo, N., Bridge, A., et al. (2013) New and continuing developments at PROSITE, *Nucleic acids research* **41**: 344-347.

Singh, A., Ma, D., and Kaplan, D.L. (2000) Enzyme-Mediated Free Radical Polymerization of Styrene, *Biomacromolecules* **1**: 592-596.

Singh, G., Bhalla, A., Kaur, P., Capalash, N., and Sharma, P. (2011) Laccase from prokaryotes: a new source for an old enzyme, *Reviews in Environmental Science and Bio/Technology* **10**: 309-326.

Sirim, D., Wagner, F., Wang, L., Schmid, R.D., and Pleiss, J. (2011) The Laccase Engineering Database: a classification and analysis system for laccases and related multicopper oxidases, *Database (Oxford)* 2011: bar006.

Skandalis, A., Encell, L.P., and Loeb, L.A. (1997) Creating novel enzymes by applied molecular evolution, *Chemistry & biology* **4**: 889-898.

Ślesak, I., Ślesak, H., and Kruk, J. (2017) RubisCO Early Oxygenase Activity: A Kinetic and Evolutionary Perspective, *BioEssays* **39**: 1700071.

Smith, M.A., and Arnold, F.H. (2014) Designing libraries of chimeric proteins using SCHEMA recombination and RASPP. In: *Directed Evolution Library Creation*: Springer. 335-343.

Smith, S.A., and Tabita, F.R. (2003) Positive and Negative Selection of Mutant Forms of Prokaryotic (Cyanobacterial) Ribulose-1,5-bisphosphate Carboxylase/Oxygenase, *Journal of Molecular Biology* **331**: 557-569.

Solomon, E.I., Sundaram, U.M., and Machonkin, T.E. (1996) Multicopper Oxidases and Oxygenases, *Chemical Reviews* **96**: 2563-2606.

Sonnhammer, E.L., Eddy, S.R., Birney, E., Bateman, A., and Durbin, R. (1998) Pfam: multiple sequence alignments and HMM-profiles of protein domains, *Nucleic acids research* **26**: 320-322.

Sperling, E.A., Frieder, C.A., Raman, A.V., Girguis, P.R., Levin, L.A., and Knoll, A.H. (2013) Oxygen, ecology, and the Cambrian radiation of animals, *Proceedings of the National Academy of Sciences of the United States of America* **110**: 13446-13451.

Steipe, B., Schiller, B., Pluckthun, A., and Steinbacher, S. (1994) Sequence statistics reliably predict stabilizing mutations in a protein domain, *Journal of Molecular Biology* **240**: 188-192.

Stemmer, W.P. (1994) Rapid evolution of a protein *in vitro* by DNA shuffling, *Nature* **370**: 389-391.

Sulpice, R., Tschoep, H., Von Korff, M., Buessis, D., Usadel, B., Hoehne, M., et al. (2007) Description and applications of a rapid and sensitive non-radioactive microplate-based assay for maximum and initial activity of D-ribulose-1, 5-bisphosphate carboxylase/oxygenase, *Plant, cell & environment* **30**: 1163-1175.

Sullivan, B.J., Durani, V., and Magliery, T.J. (2011) Triosephosphate isomerase by consensus design: dramatic differences in physical properties and activity of related variants, *Journal of molecular biology* **413**: 195-208.

Sullivan, B.J., Nguyen, T., Durani, V., Mathur, D., Rojas, S., Thomas, M., et al. (2012) Stabilizing proteins from sequence statistics: the interplay of conservation and correlation in triosephosphate isomerase stability, *Journal of molecular biology* **420**: 384-399.

Sun, Z., Lonsdale, R., Kong, X.D., Xu, J.H., Zhou, J., and Reetz, M.T. (2015) Reshaping an enzyme binding pocket for enhanced and inverted stereoselectivity: use of smallest amino acid alphabets in directed evolution, *Angewandte Chemie* **54**: 12410-12415.

Swofford, D. (1984) PAUP: Phylogenetic Analysis Using Parsimony. Version 2.3. Privately printed documentation, *Illinois Natural History Survey, Champaign, Illinois*.

Tabita, F.R. (1999) Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A different perspective, *Photosynthesis Research* **60**: 1-28.

Tabita, F.R., Hanson, T.E., Li, H., Satagopan, S., Singh, J., and Chan, S. (2007) Function, structure, and evolution of the RubisCO-like proteins and their RubisCO homologs, *Microbiology and molecular biology reviews : MMBR* **71**: 576-599.

Tabita, F.R., Satagopan, S., Scott, S.S., Kreel, N.E., and Hanson, T.E. (2008) Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships, *Journal of Experimental Botany* **59**: 1515-1524.

Tadesse, M.A., D'Annibale, A., Galli, C., Gentili, P., and Sergi, F. (2008) An assessment of the relative contributions of redox and steric issues to laccase specificity towards putative substrates, *Organic & Biomolecular Chemistry* **6**: 868-878.

Taylor, T.C., and Andersson, I. (1996) Structural transitions during activation and ligand binding in hexadecameric Rubisco inferred from the crystal structure of the activated unliganded spinach enzyme, *Nature Structural & Molecular Biology* **3**: 95-101.

Taylor, T.C., and Andersson, I. (1997) The structure of the complex between rubisco and its natural substrate ribulose 1, 5-bisphosphate, *Journal of molecular biology* **265**: 432-444.

Tcherkez, G.G.B., Farquhar, G.D., and Andrews, T.J. (2006) Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized, *Proceedings of the National Academy of Sciences* **103**: 7246-7251.

Teixeira, D., Lalot, T., Brigodiot, M., and Marechal, E. (1999) β-Diketones as Key Compounds in Free-Radical Polymerization by Enzyme-Mediated Initiation, *Macromolecules* **32**: 70-72.

Tokuriki, N., Stricher, F., Serrano, L., and Tawfik, D.S. (2008) How Protein Stability and New Functions Trade Off, *PLOS Computational Biology* **4**: e1000002.

Torres-Salas, P., Mate, D.M., Ghazi, I., Plou, F.J., Ballesteros, A.O., and Alcalde, M. (2013) Widening the pH activity profile of a fungal laccase by directed evolution, *ChemBioChem* **14**: 934-937.

Toscano, M.D., De Maria, L., Lobedanz, S., and Østergaard, L.H. (2013) Optimization of a Small Laccase by Active-Site Redesign, *ChemBioChem* **14**: 1209-1211.

Tsujimoto, T., Uyama, H., and Kobayashi, S. (2001) Polymerization of vinyl monomers using oxidase catalysts, *Macromolecular Bioscience* **1**: 228-232.

UniProt Consortium (2019) UniProt: a worldwide hub of protein knowledge, *Nucleic acids research* **47**: 506-515.

Vicente, A.I., Viña-Gonzalez, J., Mateljak, I., Monza, E., Lucas, F., Guallar, V. and Alcalde, M. (2019). Enhancing Thermostability by Modifying Flexible Surface Loops in an Evolved High-Redox Potential Laccase. *AiChE journal. American Institute of Chemical Engineers*. In press.

Villarroya, S., Thurecht, K.J., and Howdle, S.M. (2008) HRP-mediated inverse emulsion polymerisation of acrylamide in supercritical carbon dioxide, *Green Chemistry* **10**: 863-867.

Viña-Gonzalez, J., Gonzalez-Perez, D., Ferreira, P., Martinez, A.T., and Alcalde, M. (2015) Focused Directed Evolution of Aryl-Alcohol Oxidase in *Saccharomyces cerevisiae* by Using Chimeric Signal Peptides, *Applied and environmental microbiology* **81**: 6451-6462.

Viña-Gonzalez, J., Jimenez-Lalana, D., Sancho, F., Serrano, A., Martinez, A.T., Guallar, V., and Alcalde, M. (2019) Structure-Guided Evolution of Aryl Alcohol Oxidase from *Pleurotus eryngii* for the Selective Oxidation of Secondary Benzyl Alcohols, *Advanced Synthesis & Catalysis*.**361**: 2514-2525.

Virk, A.P., Sharma, P., and Capalash, N. (2012) Use of laccase in pulp and paper industry, *Biotechnology Progress* **28**: 21-32.

Viswanath, B., Rajesh, B., Janardhan, A., Kumar, A.P., and Narasimha, G. (2014) Fungal laccases and their applications in bioremediation, *Enzyme research* **2014**: 163242.

Walker, J.C. (1983) Possible limits on the composition of the Archaean ocean, *Nature* **302**: 518-520.

Wang, Q., Buckle, A.M., Foster, N.W., Johnson, C.M., and Fersht, A.R. (1999) Design of highly stable functional GroEL minichaperones, *Protein science : a publication of the Protein Society* **8**: 2186-2193.

Wang, Q., Buckle, A.M., and Fersht, A.R. (2000) Stabilization of GroEL minichaperones by core and surface mutations, *Journal of molecular biology* **298**: 917-926.

Watanabe, K., Ohkuri, T., Yokobori, S.-i., and Yamagishi, A. (2006) Designing Thermostable Proteins: Ancestral Mutants of 3-Isopropylmalate Dehydrogenase Designed by using a Phylogenetic Tree, *Journal of Molecular Biology* **355**: 664-674.

Whelan, S., and Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach, *Molecular Biology and Evolution* **18**: 691-699.

Whitfield, J.H., Zhang, W.H., Herde, M.K., Clifton, B.E., Radziejewski, J., Janovjak, H., et al. (2015) Construction of a robust and sensitive arginine biosensor through ancestral protein reconstruction, *Protein Science* **24**: 1412-1422.

Whitney, S.M., and Andrews, T.J. (1998) The CO2/O2 specificity of single-subunit ribulose-bisphosphate carboxylase from the dinoflagellate, *Amphidinium carterae*, *Functional Plant Biology* **25**: 131-138.

Whitney, S.M., Houtz, R.L., and Alonso, H. (2011) Advancing our understanding and capacity to engineer nature's CO_2-sequestering enzyme, Rubisco, *Plant physiology* **155**: 27-35.

Wijma, H.J., Floor, R.J., and Janssen, D.B. (2013) Structure-and sequence-analysis inspired engineering of proteins for enhanced thermostability, *Current opinion in structural biology* **23**: 588-594.

Wilson, D., Pethica, R., Zhou, Y., Talbot, C., Vogel, C., Madera, M., et al. (2009) SUPERFAMILY--sophisticated comparative genomics, data mining, visualization and phylogeny, *Nucleic acids research* **37**: 380-386.

Wilson, R.H., Alonso, H., and Whitney, S.M. (2016) Evolving *Methanococcoides burtonii* archaeal Rubisco for improved photosynthesis and plant growth, *Scientific Reports* **6**: 22284.

Wilson, R.H., and Whitney, S.M. (2017) Improving CO_2 Fixation by Enhancing Rubisco Performance. In: *Directed Enzyme Evolution: Advances and Applications*. Alcalde, M. (ed). Cham: Springer International Publishing. 101-126.

Williams, G., Nelson, A., and Berry, A. (2004) Directed evolution of enzymes for biocatalysis and the life sciences, *Cellular and molecular life sciences CMLS* **61**: 3034-3046.

Wirtz, P., and Steipe, B. (1999) Intrabody construction and expression III: engineering hyperstable V(H) domains, *Protein science : a publication of the Protein Society* **8**: 2245-2250.

Xu, F. (1997) Effects of redox potential and hydroxide inhibition on the pH activity profile of fungal laccases, *The Journal of Biological Chemistry* **272**: 924-928.

Xu, F., Palmer, A.E., Yaver, D.S., Berka, R.M., Gambetta, G.A., Brown, S.H., and Solomon, E.I. (1999) Targeted Mutations in a *Trametes villosa* Laccase axial perturbations of the T1 copper, *The Journal of Biological Chemistry* **274**: 12372-12375.

Yamashiro, K., Yokobori, S., Koikeda, S., and Yamagishi, A. (2010) Improvement of *Bacillus circulans* beta-amylase activity attained using the ancestral mutation method, *Protein Engineering, Design and Selection* **23**: 519-528.

Yang, H., Sun, H., Zhang, S., Wu, B., and Pan, B. (2015) Potential of acetylacetone as a mediator for *Trametes versicolor* laccase in enzymatic transformation of organic pollutants, *Environmental Science and Pollution Research* **22**: 10882-10889.

Yang, Z. (1997) PAML: a program package for phylogenetic analysis by maximum likelihood, *Bioinformatics* **13**: 555-556.

Yang, Z., Kumar, S., and Nei, M. (1995) A new method of inference of ancestral nucleotide and amino acid sequences, *Genetics* **141**: 1641-1650.

Yang, Z., and Nielsen, R. (2000) Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models, *Molecular Biology and Evolution* **17**: 32-43.

Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A.M. (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites, *Genetics* **155**: 431-449.

Yokoyama, S., and Radlwimmer, F.B. (2001) The molecular genetics and evolution of red and green color vision in vertebrates, *Genetics* **158**: 1697-1710.

Yokoyama, S., Yang, H., and Starmer, W.T. (2008) Molecular basis of spectral tuning in the red-and green-sensitive (M/LWS) pigments in vertebrates, *Genetics* **179**: 2037-2043.

Zahnd, C., Wyler, E., Schwenk, J.M., Steiner, D., Lawrence, M.C., McKern, N.M., et al. (2007) A Designed Ankyrin Repeat Protein Evolved to Picomolar Affinity to Her2, *Journal of Molecular Biology* **369**: 1015-1028.

Zakas, P.M., Brown, H.C., Knight, K., Meeks, S.L., Spencer, H.T., Gaucher, E.A., and Doering, C.B. (2017) Enhancing the pharmaceutical properties of protein drugs by ancestral sequence reconstruction, *Nature biotechnology* **35**: 35-37.

Zhu G., Kurek I., Liu L. (2010) Engineering photosynthetic enzymes involved in CO2–assimilation by gene shuffling, In: *Advances in Photosynthesis and Respiration: The Chloroplast* Rebeiz C.A., Benning C., Bohnert H.J., Daniell H., Hoober J.K., Lochtenthaler H.K., Portis A.R., Tripathy B.C. (eds), Springer Netherlands, Dordrecht, the Netherlands, pp 307–322

Zumarraga, M., Bulter, T., Shleev, S., Polaina, J., Martinez-Arias, A., Plou, F.J., et al. (2007) *In vitro* Evolution of a Fungal Laccase in High Concentrations of Organic Cosolvents, *Chemistry & Biology* **14**: 1052-1064.

8 ANNEX

6.2.1 Publications

Molina-Espeja, P., Vina-Gonzalez, J., Gomez-Fernandez, B.J., Martin-Diaz, J., Garcia-Ruiz, E., and Alcalde, M. (2016) Beyond the outer limits of nature by directed evolution, *Biotechnology advances* **34**: 754-767.

Gomez-Fernandez, B.J., Garcia-Ruiz, E., Martin-Diaz, J., Gomez de Santos, P., Santos-Moriano, P., Plou, Ballesteros, A., Garcia, M., Rodriguez, M., Risso, V.A., Sanchez-Ruiz, J.M., Whitney, S.M. and Alcalde, M. (2018) Directed -*in vitro*- evolution of Precambrian and extant Rubiscos, *Scientific Reports* 8: 5532.

Gomez-Fernandez B.J., Risso V.A., Rueda A., Sanchez-Ruiz J.M. and Alcalde M. (2019a). Directed evolution of an ancestral laccase, Manuscript in preparation.

Gomez-Fernandez B.J., Risso V.A., Sanchez-Ruiz J.M. and Alcalde M. (2019b) Consensus design of an evolved high-redox potential laccase. *Frontiers in Bioengineering and Biotechnology,* Submitted.

6.2.2 Poster presentations

Bernardo J. Gomez-Fernandez, Andres M. Rueda, Valeria A. Risso, Jose M. Sanchez-Ruiz and Miguel Alcalde (2019). "Travelling back and forth: Directed evolution of an ancestral high-redox potential laccase". SEBBM 2019 16-19 July. Madrid (Spain).

Bernardo J Gomez-Fernandez and Miguel Alcalde (2018). "Consensus design to increase thermostability in an evolved high-redox potential laccase". II JORNADAS ESPAÑOLAS DE BIOCATALISIS 2018 25-26 Jun. Oviedo (Spain).

Bernardo J Gomez-Fernandez and Miguel Alcalde (2016). "Recombination of consensus mutations to increase thermostability in an evolved high-redox potential laccase". BIO IBEROAMERICA 2016. Jun 5-8. Salamanca (Spain)

Gonzalez-Perez, D., Molina-Espeja, P., Garcia-Ruiz, E., Mate, D.M., Viña, X., Martin, J., Gomez, B.J., Vicente A.I., Mateljak, I., Ballesteros, A.O., Alcalde, M. (2014). "Directed evolution of the ligninolityc consortium". 4th International conference on Novel Enzymes. Oct, 14-17. Gant (Belgium).

Gonzalez-Perez D., Garcia-Ruiz E., Gomez Fernandez B.J., Ruiz-Dueñas F.J., Martinez A.T., Alcalde M. (2014)."Screening mutant libraries of versatile peroxidase from *Pleurotus*

eryngii to enhance oxidative stability". XXXIV National Congress of Spanish society of Biochemistry and Molecular Biology (SEBBM-2014). Sep, 9-12. (Granada, Spain).

Eva Garcia-Ruiz, Bernardo J. Gomez-Fernandez, Francisco J. Plou, Miguel Alcalde. (2014) "Engineering approaches for directed rubisco evolution". European Cooperation in Science and Technology (COST) Training School. May 28–Jun 1. Sienna (Italy).

Martin-Diaz, J., Garcia-Ruiz, E., Molina-Espeja, P., Gonzalez-Perez, D., Gomez-Fernandez B.J., Alcalde, M. (2014)." Exploring substrate promiscuity of Unspecific Peroxygenase from *Agrocybe aegerita* by neutral genetic drift". 3rd Multistep Enzyme Catalyzed Processes Congress (MECP14). Apr 7-10. Madrid (Spain).

Santos-Moriano, P., Gmez B.J., Martin-Diaz, J., Melchor, B., Gonzalez-Perez, D., Molina, P., Viña-Gonzalez, J., Rodriguez-Colinas, B., Garcia-Ruiz, E., Mate, D.M., Fernandez-Arrojo, L., Torres-Salas, P., Plou, F.J., Alcalde, M., Ballesteros, AO. (2013). "The Applied Biocatalysis Group at CSIC: activities in the field of bioenergy". CSIC-IMDEA Energy around R&D+i in energetic technologies of Madrid community. Oct, 17. Mostoles (Spain).

Bernardo J. Gomez-Fernandez, Eva Garcia -Ruiz, Javier Martin-Diaz, Miguel Alcalde. (2013) "Directed evolution of chloroperoxidase from *Caldariomyces fumago* for functional expression in yeast." XXXVI National Congress of Spanish society of Biochemistry and Molecular Biology. Sep, 4-6. Madrid (Spain).

Eva Garcia-Ruiz, Bernardo J. Gomez-Fernandez, Francisco J. Plou, Miguel Alcalde. (2013). "Engineering approaches for directed rubisco evolution". XXXVI National Congress of Spanish society of Biochemistry and Molecular Biology. Sep, 4-6. Madrid (Spain).

B.J. Gomez, A. Prieto, V. Barba, J. Martin, and M.J. Martinez. (2012). "Interesterification of triolein with short-chain alcohols or esters for biodiesel production catalized by the *Ophiostoma piceae* sterol esterase". IV National Congress of Industrial Microbiology and Microbial Biotechnology. Nov, 14-16. Salamanca (Spain).

6.2.3 Oral communications

Bernardo J. Gomez-Fernandez, Andres M. Rueda, Valeria A. Risso, Jose M. Sanchez-Ruiz and Miguel Alcalde (2019). "Travelling back and forth: Directed evolution of an ancestral high-redox potential laccase". BIOTEC 2019 10-13 Jun. Vigo (Spain).

Bernardo J Gomez-Fernandez and Miguel Alcalde (2018). "Consensus design to increase thermostability in an evolved high-redox potential laccase". II JORNADAS ESPAÑOLAS DE BIOCATALISIS 2018 25-26 Jun. Oviedo (Spain).